ALTERED NUCLEOTIDE SEQUENCE IN CD40 LIGAND PROMOTER

This application claims priority under 35 U.S.C. §119 of U.S. application Serial No. 60/153,625, filed September 13, 1999.

FIELD OF THE INVENTION

This invention relates to autoimmune and inflammatory diseases, especially Rheumatoid Arthritis (RA). This invention also relates to diseases or conditions in which an elevated CD40 ligand (CD40L) expression is a factor. In addition, the invention relates to the causes and progression of autoimmune and inflammatory diseases, and diseases in which elevated CD40L expression is a factor, especially RA. The invention further relates to new and improved diagnostic and therapeutic methods for autoimmune and inflammatory diseases, and diseases in which elevated CD40L expression is a factor, especially RA.

BACKGROUND OF THE INVENTION

CD4+ T-lymphocytes play a central role in the regulation of immune and inflammatory responses. After antigen-specific activation of T helper (Th) cells through their antigen receptor (TCR), the highly regulated expression of CD40 ligand (CD40L; CD154) on the T cell membrane mediates activation signals to interacting CD40+ target cells, including B cells, monocytes, dendritic cells, and activated endothelial cells and fibroblasts. When the expression of CD40L is altered, deficient immune responses, as are associated with mutated CD40L, or systemic immune activation, as has been observed in association with prolonged CD40L expression, can result. Furthermore, as discussed

herein, CD40L has been implicated in autoimmune diseases, including systemic lupus erythematosus (SLE) and RA.

CD40L in the immune response

Human CD40L, a type II transmembrane glycoprotein of 33 kD, belongs to the tumor necrosis factor (TNF) superfamily of cell surface interaction molecules (1). Engagement of CD40 by Th cell surface CD40L provides an essential signal for B cell activation and also mediates activation of macrophages, dendritic cells, endothelial cells and synovial cells (2-20). Mutations in the CD40L gene are responsible for the immunodeficiency of X chromosome-linked hyper-IgM syndrome (21-27), and CD40L has been found to play a critical role in systemic autoimmune diseases, including SLE and RA (28-35). Moreover, elevated levels of CD40L may play a role in other diseases and conditions such as atherosclerosis and transplant rejections.

CD40L is predominantly expressed on the CD4+ Th cell subset, although some CD8+ T cells, basophils, pulmonary mast cells, platelets, and activated B cells, have also been described as CD40L+ (1,2,28,29,32,33,36-39). *In vivo*, CD40L expression is mostly restricted to secondary lymphoid follicles, a site of immunoglobulin (Ig) class switching (40,41). Activation of T cells through the TCR-CD3 complex and CD28 results in rapid induction of T cell surface CD40L, with peak expression observed at 6 hours and markedly decreased expression by 24 hours (28). The molecular structure of CD40L provides a potential site for proteolytic cleavage and shedding from the T cell surface, and a soluble form of the molecule has been reported (4,42-47). The rapid on and off of the T cell surface expression of CD40L following antigen-specific Th cell activation is a central point of regulation of the humoral immune response to T-dependent antigens. CD40L is primarily responsible for linked recognition of antigen by T and B cells, and it is the tight control of its expression that assures the fairly restricted specificity of Th-dependent antibody responses.

CD40L:CD40 interaction is required for the differentiation of B cells to IgG, IgA, and IgE production (6,40,48-56). The Ig classes secreted by B cells activated through CD40 are modulated by the dominant cytokine (IL-2, IL-4, or TGF-b) to which those B cells are exposed (6,8,49-52,56). CD40 ligation can also induce B cell secretion of TNF-

a, IL-6, TGFβ, or IL-10, which may further drive B cell differentiation and promote inflammation (56,57). The essential role of CD40L in human B cell function, particularly Ig class switching to mature isotypes, is confirmed by the demonstration that abnormal CD40L, based on any of a number of point mutations or deletions in the X chromosome gene encoding that molecule, is the molecular basis of the X-linked hyper-IgM syndrome (21-27). This immunodeficiency syndrome is characterized by high levels of serum IgM, low or absent levels of IgG and IgA, and absent IgE. Its clinical features include recurrent infections and an increased incidence of lymphoma. The significance of CD40L:CD40 interaction for induction of Ig class switch recombination and in the generation of a mature immune response has also been confirmed in CD40L knock-out mice (40,41).

Beyond its important role in Ig class switching, CD40L promotes B cell antigen presentation function, clonal expansion, and rescue from apoptosis in the germinal center. Ligation of B cell CD40 by CD40L leads to the formation of homotypic adhesions among B cells, heterotypic adhesions between T and B cells, and augmented expression of several B cell surface activation, adhesion and co-stimulatory molecules, including CD23, CD54 (ICAM-1), CD80 (B7-1), CD86 (B7-2), MHC class II, and CD95 (Fas) (1,10,11,58-62). Particularly as a result of increased CD80 and 86, high density B cells activated through CD40 have augmented capacity to stimulate T cell activation and proliferation (11). CD40L-expressing T cells can rescue B cells from apoptosis following surface Ig ligation by antigen or promote apoptosis through the Fas pathway in the absence of B cell receptor signals (48,59-62). Thus, as long as specific antigen is available for triggering of T and B cell antigen receptors, CD40L:CD40 interaction has the potential to promote and perpetuate T cell-B cell interaction, with concomitant cytokine production, antibody production, and determinant spreading of the antibody response to a widening range of antigenic specificities.

However, the functional importance of CD40L extends beyond Th-dependent antibody responses. CD40L has been implicated in macrophage and dendritic cell secretion of nitric oxide, TNFα, and IL-12 (15,20), endothelial cell activation and expression of adhesion molecules and coagulation factors (17-19), and induction of cell surface adhesion molecules and metalloproteinase enzymes by fibroblasts and

synoviocytes from RA joint tissue (16,30,34,35). The capacity of Th cell CD40L to mediate induction of effector functions by a wide range of CD40+ target cells may be particularly significant in a localized inflammatory setting, such as the RA joint, where all of these cell types are chronically gathered together in an anatomically confined space.

Regulation of CD40L expression

As noted, the normally brief expression of CD40L after TCR-mediated Th cell activation reflects the important role of that molecule in the maintenance of the fine specificity of an immune response. Prolonged or ectopic expression of CD40L, as observed in SLE (28,29,32), may contribute to polyclonal B cell activation and the induction of undesired antibody specificities, as well as cytokine production by macrophages and dendritic cells, and endothelial and synoviocyte activation. Several mechanisms are known or postulated to control the expression of CD40L: a) transcriptional regulation; b) post-transcriptional regulation of CD40L mRNA; c) release of CD40L protein onto the cell membrane; and d) enzymatic release of soluble CD40L (sCD40L) from the cell membrane or from intracellular stores (1,4,28,43-45,63). The close correlation of both kinetics and quantity of CD40L mRNA expression with cell surface sCD40L protein expression suggests the importance of transcriptional regulation (47,64). The functional significance of transcriptional or posttranscriptional CD40L controls is demonstrated by the recent report of a 4-5 fold increase in production of a T-dependent antibody when CD40L mRNA and protein were increased by less than 2 fold (64).

The genomic structure of human CD40L has been characterized (4,65-68). The gene is located on chromosome Xq26-27 and includes five exons and four introns, a 3' untranslated region that contains a polymorphic (CA)n/(GT)n repeat, and a 5' promoter. The approximately 500 bp 5' of the transcription initiation site have been shown to contain the key regulatory elements that confer transcription in a transfection system, although one abstract has suggested that a motif in a 3' enhancer region magnifies the level of transcription (63,69,70). An unusual feature of the 5' promoter region is a poly-A tract, more commonly seen in the 3' segment of genes. More typical features of the promoter include a TATA-like sequence and two NF-AT-binding motifs that are

important for transcriptional activity (63). Nuclear protein extracts from activated CD4 T cell lines bound to an oligonucleotide probe containing the proximal NF-AT element (-62 to -69 5' of the transcription initiation site) in electromobility shift assays (EMSA), supporting a role for that transcription factor in CD40L promoter function (63). The importance of other binding motifs in the proximal promoter and their interacting proteins in the initiation of human CD40L transcription has not yet been addressed.

CD40L expression in autoimmune diseases, including RA

While hyperIgM immunodeficiency syndrome is based on mutation and impaired function of CD40L, patients with systemic autoimmune diseases characterized by excessive Th cell-dependent B cell activation and differentiation have been shown to have increased or prolonged expression of CD40L. This altered expression has been best documented in SLE, the prototypic systemic autoimmune disease in which a constitutive expression of CD40L on T cells has been found in patients with clinically active disease (28,29,32). Moreover, following stimulation of SLE T cells with the non-specific activators phorbol myristate acetate (PMA) and ionomycin, cell surface expression of the usually tightly regulated CD40L is prolonged up to 48 hours in some patients (28). Concordant with those results are additional data demonstrating increased CD40L mRNA stability in patients with SLE. Increased production of CD40L is most readily discerned by quantitation of the soluble form of CD40L. While normal subjects have either undetectable or low pg/ml levels of sCD40L in sera, sera from patients with SLE have significant elevations of sCD40L that is related to the degree of disease activity (46,47). These data, which confirm the predicted role of increased Th cell function in a disease, SLE, characterized by increased spontaneous Ig class switching and production of somatically mutated autoantibodies, have provided the rationale for successful pre-clinical studies, and ongoing clinical trials, of CD40L blockade in SLE (71).

Although most clinical studies of CD40L expression and function have been performed in SLE and in conditions of allograft rejection (28,29,32,72), recent data implicate that molecule in the pathogenesis of RA as well. CD40L+ T cells have been demonstrated in a subset of RA peripheral blood, synovial fluid (SF), and synovial tissue (ST) samples (31,33). The soluble form of CD40L is also present in some RA SF (31).

The pathogenic potential of synovial CD40L is indicated by the proliferation of synovial cells and their production of TNFα when triggered through CD40 (30,34,35). The presence of excess sCD40L in the peripheral circulation, is, however, less impressive in RA than in SLE (46). One study was unable to detect serum sCD40L in the same patients who did have sCD40L present in SF (31). Data from patients with RA show increased levels in some RA sera, but at generally lower concentrations than in SLE (46). It should be noted, however, that CD40L expression is not an absolute prerequisite for inflammatory polyarthritis. An interesting recent description of a patient with hyper-IgM syndrome who also had a very destructive polyarthritis indicates that mechanisms other than high CD40L, perhaps high levels of TNFa induced through non-CD40 pathways, can produce clinical RA (73). However, the general importance of CD40L:CD40 interactions in inflammatory arthritis syndromes is documented by abrogation of disease by the specific blockade of CD40L in a common model of inflammatory arthritis, collagen-induced arthritis (74). Characterization of the DNA elements and transcription factors that mediate CD40L expression in the RA synovium is a primary aim of the proposed research.

Susceptibility genes in RA

Studies of the genetic basis of susceptibility to RA and of disease severity have focused on the HLA-DR locus, with DRB1*01 and 04 conferring increased risk. Examination of non-HLA susceptibility genes in RA is at an early stage. It is of interest, however, that a microsatellite near CD40L has suggested linkage to RA in patients who are DR4-/DR1- (75). In those subjects, the expression of the (GT)21 allele, located in the 3' untranslated region of the CD40L gene, increased the relative risk of acquiring RA more than 11 times, but was more important in males than females. A second recent study reports that the maximum lod score (MLS) for a site near the CD40L gene on chromosome X, between markers DXS1227 and DXS1200, was 2.93 and a region 2 mM to the right of DXS1232 had a MLS of 3.03 (76). However, additional study is needed to determine the relationship of these observations, if any, to CD40L regulation and function.

Microsatellite instability

Genome fidelity is a high priority in biologic systems. With each replication of DNA during cell division, the potential for error, resulting in point mutations, deletions, or additions to the genome, is offset by the complex machinery of DNA repair. Somatic hypermutation of Ig genes is an exception to this generalization, permitting the controlled mutation of a limited stretch of DNA that spans the 5' region of heavy and light chain Ig genes when an antigen-activated B cell receives the appropriate complement of Th-derived signals (77). Other than this very specialized function of B lymphocytes, somatic mutation has primarily been observed in the setting of malignancies. Microsatellite instability is a type of mutational event that refers to variability in the size of nucleotide repeats and has been associated with tumors of the replication error (RER) phenotype, such as familial colorectal cancer (78-87). A growing number of examples of poly-A nucleotide tracts that are associated with microsatellite instability and malignancy has drawn attention to these motifs and the role that they may play in induction of mutation. A poly-A tract in the coding sequence of the human MSH2 mismatch repair gene is one of these unstable microsatellites (83,88), and a poly-A repeat in exon 3 of the transforming growth factor (TGF) bII receptor gene is also subject to mutations that can result in a frameshift in RER tumors (78,84,86,87). The general rule is that repetitive sequences are copied with less fidelity than nonrepetitive sequences, with additions or deletions in those repetitive sequences sometimes resulting in a "mutator phenotype" (89). While many of the altered microsatellite sequences have no effect on the function of the organism, others impact important cellular pathways (90).

Several examples of poly-A tracts in gene promoters suggest that microsatellite instability may also affect promoter function (91-94). A 13 bp poly-A tract in the mammalian ME1 gene promoter binds a transcription factor, MBPa; is associated with DNA bending; and can initiate transcription from the mid-region of the poly-A tract (93,94). It has been suggested that the DNA bending in this region may be important in permitting DNA polymerase entry into the region of transcription initiation (95). A recent abstract reports on a poly-A tract of variable length in the LTR promoter region of the HRES-1 human endogenous retroviral sequence on chromosome 1q41 that is associated

with nearby mutations, although promoter function was not investigated in that study (96).

Alterations in nucleotides in a promoter region can change the conformation of the promoter itself, including its regulatory elements, affect the binding of transcription factors, and up- or down-regulate gene expression. Thus there is a need in the art to further explore the factors which influence CD40L expression, especially the promoter region, to better understand the causes and progression of autoimmune and inflammatory diseases, and diseases or conditions in which elevated CD40L expression is a factor, especially RA. Such knowledge may lead to the discovery for new and improved treatment methods for RA and other autoimmune and/or inflammatory diseases. There is further a need in the art for better diagnostic procedures to evaluate diseases or conditions in which elevated CD40L expression is a factor, as well as identify individuals who are at a risk of contracting such diseases. The present invention addresses these and other needs in the art.

SUMMARY OF THE INVENTION

The present invention provides an altered CD40L promoter, and uses of the altered CD40L promoter in the study, diagnosis, and treatment of a variety of inflammatory and autoimmune diseases, as well as diseases in which elevated expression of CD40L is a factor, especially Rheumatoid Arthritis (RA). Applicants have surprisingly discovered that the altered promoter is increased in prevalence in individuals with RA. Without being bound to any specific theory, it is believed that this altered promoter contributes to increased gene expression, protein production, and inflammation in the synovial membrane. The altered promoter sequence and related proteins, such as, *e.g.*, transcriptional factors, which interacts with the altered CD40L promoter therefore present new therapeutic targets for the diagnosis and treatment of a variety of diseases, especially RA.

Thus, nucleic acids corresponding to the altered promoter sequence or parts thereof; proteins peptides, or other factors which interact with the altered CD40L promoter sequence, antibodies to the altered CD40L promoter sequence, and cells transformed with nucleic acids containing the altered promoter sequence, as well as transgenic animals comprising such nucleic acids, that possess various utilities, are described herein for the diagnosis, therapy and continued investigation of diseases and conditions in which an elevated expression of CD40L is a factor, especially RA.

The invention provides a method for detecting an alteration of the CD40L promoter sequence associated with inflammatory and autoimmune disorders, or another disorder in which elevated CD40L expression is a factor, especially RA, comprising obtaining a nucleic acid sample from an individual at risk for, diagnosed with, or suspected of having, RA or another inflammatory or autoimmune disease, and sequencing the CD40L promoter sequence from said sample. In particular, such methods can identify normal human alleles as well as altered alleles of the CD40L promoter which are causative of or contribute to such disorders, especially RA.

The invention also invention provides a method for identifying individuals predisposed to or having an inflammatory or autoimmune disease, or a disease in which elevated CD40L expression is a factor, such as RA, comprising obtaining a nucleic acid sample from an individual diagnosed with, suspected of having, or at risk for, such a disease, and sequencing the CD40L promoter.

The invention also provides a method for identifying individuals predisposed to or having a an inflammatory or autoimmune disease such as RA, or another disease in which elevated CD40L expression is a factor, comprising obtaining cells that contain nucleic acid comprising the CD40L promoter, and under non-pathological conditions, measuring a level of transcriptional activity of the nucleic acid encoding for CD40L.

The invention further provides a method for identifying individuals predisposed to or having an inflammatory or autoimmune disease, especially RA, or a related disorder, comprising obtaining cells from an individual that express nucleic acid encoding CD40L, and measuring CD40L transcriptional activity. Alternatively, CD40L could be isolated from that individual to investigate, for example, whether CD40L mRNA transcription or CD40L expression levels differ from typical levels.

The invention also provides a method for identifying putative agents that affect an inflammatory or autoimmune disease, or a disease or condition in which elevated CD40L levels is a factor, especially RA, comprising adding one or more of said agents to a reconstituted system comprising the altered promoter sequence and all or parts of the CD40L gene, and detecting a change in CD40L transcriptional activity.

The invention also provides a method for identifying putative agents that affect an inflammatory or autoimmune disease, or a disease or condition in which elevated CD40L expression is a factor, especially RA, comprising adding one or more said agents, such as a transcription factor, to the altered promoter sequence, and detecting a conformational change in the promoter sequence.

The invention also provides cellular models of inflammatory or autoimmune diseases such as RA, or related disorders, that comprise the altered promoter sequence and all or part of the CD40L gene, which can be used as a therapeutic target for the development of drugs that interact with the altered promoter sequence, and thus can useful in the treatment and prevention of these disorders.

Further the invention provides for a method for identifying substances that modulate CD40L transcriptional activity, comprising contacting a sample containing one or more substances with the reconstituted or cellular model comprising the altered promoter sequence or fragments thereof, measuring CD40L transcription, and determining whether a change in CD40L transcriptional activity occurs. In a preferred embodiment, the substance is a negative regulatory element, *i.e.*, downregulates CD40L transcriptional activity. In another preferred embodiment, the substance is a positive regulatory element, *i.e.*, stimulates CD40L transcriptional activity.

These and other aspects of the invention are further elaborated in the Detailed Description of the Invention and Examples, infra.

DESCRIPTION OF THE DRAWINGS

FIG. 1 shows that protein complexes from activated peripheral blood T cells bind to oligonucleotides derived from the proximal CD40L promoter. ^{32}P-labeled double-stranded oligonucleotide fragments, corresponding to -88 to -57 bp (NM1-L) or -73 to

-41 bp (NM1-P) of the proximal CD40L promoter, were incubated with PBMC nuclear extracts from a healthy subject and then run on a polyacrylamide gel. Protein complexes bound to the oligonucleotides retarded the migration of the labeled promoter fragments. An oligonucleotide containing a known NF-AT site in the human IL-4 promoter served as a positive control.

FIG. 2 shows 5' flanking sequence alignment for wild-type **[SEQ ID NO: 1]** and altered **[SEQ ID NO: 2]** CD40L. The 5' flanking sequence of CD40L was amplified from genomic DNA from healthy subjects and from individuals with systemic autoimmune disease and sequenced.

FIG. 3 shows representative poly-A tract sequences **[SEQ ID NOS: 3-32]**, the results of direct sequencing of the CD40L proximal promoter from 5 arthritis synovial tissue and 2 control peripheral blood samples. All samples demonstrated the consensus ATT 5' of the poly-A tract and CCTTT 3' of the poly-A tract. Variability in the length of the poly-A tract was observed in all individuals studied, and the substitution of a C for an A at position -125 (indicated by *) was observed in some samples. Note that ST07, derived from a male, shows the A to C alteration in 4/4 subclones sequenced.

FIG. 4 shows a summary of CD40L promoter sequence data on patients with arthritis. ST designates synovial tissue samples; PB designates peripheral blood samples; Ethnic group designation: AS = Asian, CA = Caucasian, HI = Hispanic; "+" indicates the presence of an A to C change at position -125, corresponding to residue 331 of **SEQ ID NO: 2**. The number of subclones with an A to C change to the total number of subclones sequence is indicated.

FIG. 5 shows representative ABI Prism data demonstrating wild-type and altered poly-A tract sequence in 2 subclones from a synovial tissue sample. Genomic DNA from an RA female was amplified, subcloned, and sequenced. Two out of 7 sequenced subclones are shown. The top panel demonstrates a poly-A tract expressing the wild-type

A at position -125. The bottom panel demonstrates a poly-A tract expressing the altered A to C at position -125.

FIG. 6 shows a summary of CD40L promoter sequence data on patients with SLE or healthy subjects.

FIG. 7 shows that A to C substitution at position -125 of the CD40L proximal promoter confers increased promoter activity. CD40L promoter segments containing either A (wild-type) or C (altered) at position -125 were tested for activity by the luciferase reporter assay.

FIG. 8 a comparison between human **[SEQ ID NO: 1]** and mouse (Genbank Accession No. L47983 **[SEQ ID NO: 37])** CD40L proximal promoter sequence. Divergent nucleotides are indicated with an * and gaps in nucleotides are indicated with a –. The nucleotide positions, in relation to the transcription start site, are labeled in reference to the human sequence. The TATA box (-140 to -136), the CRE BPI-binding consensus site (-109 to -102, and the NF-AT-binding motif (-68 to -63) are underlined.

FIG. 9 shows prolonged CD40L mRNA expression in SLE PBMC compared with healthy control PBMC .

FIG. 10 shows constructs used in transient transfection and luciferase reporter assay to assess CD40L promoter activity.

FIG. 11 shows the amount of soluble CD40L in sera from patients with systemic autoimmune disease. Serum samples were collected from healthy subjects, SLE patients, and patients with other autoimmune or inflammatory conditions, including RA, systemic vasculitis, anti-phospholipid antibody syndrome, Lyme disease, and other disorders). Soluble CD40L was quantified by ELISA.

FIG. 12 shows a strategy to screen for the A to C alteration in the CD40L proximal promoter by the ARMS method, via a two-stage amplification of the poly-A tract.

FIG. 13 shows the results of an ARMS screening experiment wherein the A to C alteration was found in two patients; ST28 and ST30.

FIG. 14 shows a summary of CD40L promoter A to C alteration data in arthritis patients compared to healthy controls. Arthritis patients (termed RA but also including several OA, OA/RA, JRA, and an AVN patient) had a statistically significantly increased occurrence of C at position -125 when compared to healthy controls, with chi-square = 7.8, p=0.008.

DETAILED DESCRIPTION OF THE INVENTION

The present invention concerns altered promoter sequences for a CD40 ligand (CD40L) implicated in various diseases, including inflammatory and autoimmune diseases, especially Rheumatoid Arthritis (RA), and methods of use thereof. In particular, the invention concerns the discovery of an A to C substitution in the proximal promoter of CD40L in RA patients, which provides for new strategies to study the mechanisms of RA, new methods for RA diagnosis, and new targeted therapy to modulate CD40L expression in RA and other autoimmune diseases.

The prevalence of the altered nucleotide sequence in the proximal promoter region of CD40L is increased in genomic DNA samples isolated from RA synovial tissue and peripheral blood. Further, transcriptional activity differs between wild-type and altered CD40L promoter fragments. The altered nucleotide sequence is centered in a poly-adenine (poly-A) tract, a DNA motif that is unusual in 5' regulatory regions and that is of varying length among the sequences studied. Characterization of these CD40L DNA alterations and the transcriptional regulatory proteins that bind to the altered promoter will provide new information on the effects of genetic variability in a key immunoregulatory molecule. It is believed that, without being bound to any specific theory, the demonstration of altered promoter function and increased CD40L protein

expression in association with the altered promoter nucleotide sequence shows a pathogenic role of the altered promoter sequence in diseases or disorders where CD40L is implicated, especially RA. Based upon the invention, new targeted therapies can be developed to modulate CD40L expression and immune system activity in RA and other systemic autoimmune or inflammatory diseases, as well as related disorders. Moreover, the invention provides diagnostic methods to identify, confirm, and/or evaluate inflammatory or autoimmune diseases, and diseases or conditions in which an elevated expression of CD40L is a factor, especially RA.

Definitions

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, *e.g.*, Sambrook, 1989; Glover, 1985; Hames and Higgins, 1985; Hames and Higgins, 1984; Freshney, 1986; Perbal, 1984; Ausubel *et al.*, 1994 (116-122).

If appearing herein, the following terms shall have the definitions set out below.

As used herein, "about" or "approximately" shall mean within 50 percent, preferably within 20 percent, more preferably within 5 percent, of a given value or range.

A value which is "substantially different" from another value can mean that there is a statistically significant difference between the two values. Any suitable statistical method known in the art can be used to evaluate whether differences are significant or not. A "statistically significant" difference means a significance is determined at a confidence interval of at least 90%, more preferably at a 95% confidence interval.

"DNA" (deoxyribonucleic acid) means any chain or sequence of the chemical building blocks adenine (A), guanine (G), cytosine (C) and thymine (T), called nucleotide bases, that are linked together on a deoxyribose sugar backbone. DNA can have one strand of nucleotide bases, or two complimentary strands which may form a double helix structure.

"RNA" (ribonucleic acid) means any chain or sequence of the chemical building blocks adenine (A), guanine (G), cytosine (C) and uracil (U), called nucleotide bases, that

are linked together on a ribose sugar backbone. RNA typically has one strand of nucleotide bases.

A "polynucleotide" or "nucleotide sequence" is a series of nucleotide bases (also called "nucleotides") in DNA and RNA, and means any chain of two or more nucleotides. A nucleotide sequence typically carries genetic information, including the information used by cellular machinery to make proteins and enzymes. These terms include double or single stranded genomic and cDNA, RNA, any synthetic and genetically manipulated polynucleotide, and both sense and anti-sense polynucleotide (although only sense stands are being represented herein). This includes single- and double-stranded molecules, i.e., DNA-DNA, DNA-RNA and RNA-RNA hybrids, as well as "protein nucleic acids" (PNA) formed by conjugating bases to an amino acid backbone. This also includes nucleic acids containing modified bases, for example thio-uracil, thio-guanine and fluoro-uracil.

The polynucleotides herein may be flanked by natural regulatory sequences, or may be associated with heterologous sequences, including promoters, enhancers, response elements, signal sequences, polyadenylation sequences, introns, 5'- and 3'- non-coding regions, and the like. The nucleic acids may also be modified by many means known in the art. Non-limiting examples of such modifications include methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, and internucleotide modifications such as, for example, those with uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoroamidates, carbamates, etc.) and with charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.). Polynucleotides may contain one or more additional covalently linked moieties, such as, for example, proteins (e.g., nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), intercalators (e.g., acridine, psoralen, etc.), chelators (e.g., metals, radioactive metals, iron, oxidative metals, etc.), and alkylators. The polynucleotides may be derivatized by formation of a methyl or ethyl phosphotriester or an alkyl phosphoramidate linkage. Furthermore, the polynucleotides herein may also be modified with a label capable of providing a detectable signal, either directly or indirectly. Exemplary labels include radioisotopes, fluorescent molecules, biotin, and the like.

A "codon" is a triplet of nucleotides corresponding to an amino acid. Each amino acid is represented in DNA or RNA by one or more codons. The genetic code has some redundancy, also called degeneracy, meaning that most amino acids have more than one corresponding codon. For example, the amino acid lysine (Lys) can be coded by the nucleotide triplet or codon AAA or by the codon AAG.

The "reading frame" describes the way that a nucleotide sequence is grouped into codons. Because the nucleotides in DNA and RNA sequences are read in groups of three for protein production, it is important to begin reading the sequence at the correct amino acid, so that the correct triplets are read.

A "coding sequence" or a sequence "encoding" a polypeptide, protein or enzyme is a nucleotide sequence that, when expressed, results in the production of that polypeptide, protein or enzyme, i.e., the nucleotide sequence encodes an amino acid sequence for that polypeptide, protein or enzyme. A coding sequence is "under the control" of transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then trans-RNA spliced and translated into the protein encoded by the coding sequence. Preferably, the coding sequence is a double-stranded DNA sequence which is transcribed and translated into a polypeptide in a cell in vitro or in vivo when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxyl) terminus. A coding sequence can include, but is not limited to, prokaryotic sequences, cDNA from eukaryotic mRNA, genomic DNA sequences from eukaryotic (e.g., mammalian) DNA, and even synthetic DNA sequences. If the coding sequence is intended for expression in a eukaryotic cell, a polyadenylation signal and transcription termination sequence will usually be located 3' to the coding sequence.

The term "gene", also called a "structural gene" means a DNA sequence that codes for or corresponds to a particular sequence of amino acids which comprise all or part of one or more proteins or enzymes, and may or may not include regulatory DNA sequences, such as promoter sequences, which determine for example the conditions under which the gene is expressed. Some genes, which are not structural genes, may be transcribed from DNA to RNA, but are not translated into an amino acid sequence. Other

genes may function as regulators of structural genes or as regulators of DNA transcription. A gene encoding a protein of the invention for use in an expression system, whether genomic DNA or cDNA, can be isolated from any source, particularly from a human cDNA or genomic library.

A transcriptional or translational "control sequence" is a DNA regulatory sequence, such as a promoter, enhancer, terminator, and the like, that provide for the expression of a coding sequence in a host cell.

A transcriptional or translational "control element"or "regulatory element" is an element, such as, *e.g.*, a transcription factor, that induces, stimulates, down-regulates, or affect, the transcription or translation, respectively, of a gene or polynucleotide sequence.

A "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. The promoter sequence can be bounded at its 3' terminus by the transcription initiation site and extend upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently defined for example, by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. As described above, promoter DNA is a DNA sequence which initiates, regulates, or otherwise mediates or controls the expression of the coding DNA. A promoter may be "inducible", meaning that it is influenced by the presence or amount of another compound (an "inducer"). For example, an inducible promoter includes those which initiate or increase the expression of a downstream coding sequence in the presence of a particular inducer compound. A "leaky" inducible promoter is a promoter that provides a high expression level in the presence of an inducer compound and a comparatively very low expression level, and at minimum a detectable expression level, in the absence of the inducer.

A "signal sequence" can be included at the beginning of the coding sequence of a protein to be expressed in the periplasmic space, or outside the cell. This sequence encodes a signal peptide, N-terminal to the mature polypeptide, that directs the host cell to translocate the polypeptide. The term "translocation signal sequence" is also used to refer to a signal sequence. Translocation signal sequences can be found associated with

a variety of proteins native to eukaryotes and prokaryotes, and are often functional in both types of organisms. Proteins of the invention may be further modified and improved by adding a sequence which directs the secretion of the protein outside the host cell. The addition of the signal sequence does not interfere with the folding of the secreted protein, and evidence thereof is easily tested for using techniques known in the art and depending on the protein (e.g., tests for activity of a given protein after modification).

Polynucleotides are "hybridizable" to each other when at least one strand of one polynucleotide can anneal to another polynucleotide under defined stringency conditions. Stringency of hybridization is determined, e.g., by a) the temperature at which hybridization and/or washing is performed, and b) the ionic strength and polarity (e.g., formamide) of the hybridization and washing solutions, as well as other parameters. Hybridization requires that the two polynucleotides contain substantially complementary sequences; depending on the stringency of hybridization, however, mismatches may be tolerated. Typically, hybridization of two sequences at high stringency (such as, for example, in an aqueous solution of 0.5×SSC at 65°C) requires that the sequences exhibit some high degree of complementarity over their entire sequence. Conditions of intermediate stringency (such as, for example, an aqueous solution of 2×SSC at 65°C) and low stringency (such as, for example, an aqueous solution of 2×SSC at 55°C), require correspondingly less overall complementarity between the hybridizing sequences. (1×SSC is 0.15 M sodium chloride, 0.015 M sodium citrate.) Polynucleotides that "hybridize" to the polynucleotides herein may be of any length. In one embodiment, such polynucleotides are at least 10, preferably at least 15 and most preferably at least 20 nucleotides long. In another embodiment, polynucleotides that hybridizes are of about the same length.

The term "DNA reassembly" is used when recombination occurs between identical sequences. "DNA shuffling" refers herein to a group of *in vitro* and *in vivo* methods involving recombination of nucleic acid species.

A "protein" or "polypeptide", which terms are used interchangeably herein, comprises one or more chains of chemical building blocks called amino acids that are linked together by chemical bonds called peptide bonds.

An "enzyme" means any substance, preferably composed wholly or largely of

protein, that catalyzes or promotes, more or less specifically, one or more chemical or biochemical reactions. The term "enzyme" can also refer to a catalytic polynucleotide (e.g. RNA or DNA). A "test" enzyme is a substance that is tested to determine whether it has properties of an enzyme.

A "native" or "wild-type" protein, enzyme, polynucleotide, gene, or cell, means a protein, enzyme, polynucleotide, gene, or cell that occurs in nature.

A "parent" protein, enzyme, polynucleotide, gene, or cell, is any protein, enzyme, polynucleotide, gene, or cell, from which any other protein, enzyme, polynucleotide, gene, or cell, is derived or made, using any methods, tools or techniques, and whether or not the parent is itself native or mutant. A parent polynucleotide or gene can encode for a parent protein or enzyme.

A "mutant", "altered", "variant" or "modified" protein, enzyme, polynucleotide, gene, or cell, means a protein, enzyme, polynucleotide, gene, or cell, that has been altered or derived, or is in some way different or changed, from a parent protein, enzyme, polynucleotide, gene, or cell. An alteration in a gene includes, but is not limited to, alteration the promoter region, or other regions which affect transcription, which can result in altered expression levels of a protein. A mutant or modified protein or enzyme is usually, although not necessarily, expressed from a mutant polynucleotide or gene.

A "mutation" or "alteration" means any process or mechanism resulting in a mutant protein, polynucleotide, gene, or cell. This includes any mutation in which a protein, polynucleotide, or gene sequence is altered, any protein, polynucleotide, or gene sequence arising from a mutation, any expression product (*e.g.* protein) expressed from a mutated polynucleotide or gene sequence, and any detectable change in a cell arising from such a mutation.

"Function-conservative variants" are proteins or enzymes in which a given amino acid residue has been changed without altering overall conformation and function of the protein or enzyme, including, but not limited to, replacement of an amino acid with one having similar properties (such as, for example, acidic, basic, hydrophobic, and the like). Amino acids with similar properties are well known in the art. For example, arginine, histidine and lysine are hydrophilic-basic amino acids and may be interchangeable. Similarly, isoleucine, a hydrophobic amino acid, may be replaced with leucine,

methionine or valine. Amino acids other than those indicated as conserved may differ in a protein or enzyme so that the percent protein or amino acid sequence similarity between any two proteins of similar function may vary and may be, for example, from 70% to 99% as determined according to an alignment scheme such as by the Cluster Method, wherein similarity is based on the MEGALIGN algorithm. A "function-conservative variant" also includes a polypeptide or enzyme which has at least 60% amino acid identity as determined by BLAST or FASTA algorithms, preferably at least 75%, most preferably at least 85%, and even more preferably at least 90%, and which has the same or substantially similar properties or functions as the native or parent protein or enzyme to which it is compared.A "luminescent" substance means any substance which produces detectable electromagnetic radiation, or a change in electromagnetic radiation, most notably visible light, by any mechanism, including color change, UV absorbance, fluorescence and phosphorescence. Preferably, a luminescent substance according to the invention produces a detectable color, fluorescence or UV absorbance.

The term "host cell" means any cell of any organism that is selected, modified, transformed, grown, or used or manipulated in any way, for the production of a substance by the cell, for example the expression by the cell of a gene, a DNA or RNA sequence, a protein or an enzyme.

The term "expression system" means a host cell and compatible vector under suitable conditions, e.g. for the expression of a protein coded for by foreign DNA carried by the vector and introduced to the host cell. Common expression systems include bacteria (e.g. E. coli and B. subtilis) or yeast (e.g. S. cerevisiae) host cells and plasmid vectors, and insect host cells and Baculovirus vectors. As used herein, a "facile expression system" means any expression system that is foreign or heterologous to a selected polynucleotide or polypeptide, and which employs host cells that can be grown or maintained more advantageously than cells that are native or heterologous to the selected polynucleotide or polypeptide, or which can produce the polypeptide more efficiently or in higher yield. For example, the use of robust prokaryotic cells to express a protein of eukaryotic origin would be a facile expression system. Preferred facile expression systems include E. coli, B. subtilis and S. cerevisiae host cells and any suitable vector.

The term "transformation" means the introduction of a "foreign" (i.e. extrinsic or extracellular) gene, DNA or RNA sequence to a host cell, so that the host cell will express the introduced gene or sequence to produce a desired substance, typically a protein or enzyme coded by the introduced gene or sequence. The introduced gene or sequence may also be called a "cloned" or "foreign" gene or sequence, may include regulatory or control sequences, such as start, stop, promoter, signal, secretion, or other sequences used by a cell's genetic machinery. The gene or sequence may include nonfunctional sequences or sequences with no known function. A host cell that receives and expresses introduced DNA or RNA has been "transformed" and is a "transformant" or a "clone." The DNA or RNA introduced to a host cell can come from any source, including cells of the same genus or species as the host cell, or cells of a different genus or species.

The terms "vector", "cloning vector" and "expression vector" mean the vehicle by which a DNA or RNA sequence (e.g. a foreign gene) can be introduced into a host cell, so as to transform the host and promote expression (e.g. transcription and translation) of the introduced sequence.

Vectors typically comprise the DNA of a transmissible agent, into which foreign DNA is inserted. A common way to insert one segment of DNA into another segment of DNA involves the use of enzymes called restriction enzymes that cleave DNA at specific sites (specific groups of nucleotides) called restriction sites. Generally, foreign DNA is inserted at one or more restriction sites of the vector DNA, and then is carried by the vector into a host cell along with the transmissible vector DNA. A segment or sequence of DNA having inserted or added DNA, such as an expression vector, can also be called a "DNA construct."

A common type of vector is a "plasmid", which generally is a self-contained molecule of double-stranded DNA, that can readily accept additional (foreign) DNA and which can readily introduced into a suitable host cell. A plasmid vector often contains coding DNA and promoter DNA and has one or more restriction sites suitable for inserting foreign DNA. Promoter DNA and coding DNA may be from the same gene or from different genes, and may be from the same or different organisms. A large number of vectors, including plasmid and fungal vectors, have been described for replication

and/or expression in a variety of eukaryotic and prokaryotic hosts. Non-limiting examples include pKK plasmids (Clonetech), pUC plasmids, pET plasmids (Novagen, Inc., Madison, WI), pRSET or pREP plasmids (Invitrogen, San Diego, CA), or pMAL plasmids (New England Biolabs, Beverly, MA), and many appropriate host cells, using methods disclosed or cited herein or otherwise known to those skilled in the relevant art. Recombinant cloning vectors will often include one or more replication systems for cloning or expression, one or more markers for selection in the host, e.g. antibiotic resistance, and one or more expression cassettes. Preferred vectors include without limitations pGL-2, pcWori, pET-26b(+), pXTD14, pYEX-S1, pMAL, and pET22-b(+). Other vectors may be employed as desired by one skilled in the art. Routine experimentation in biotechnology can be used to determine which vectors are best suited for used with the invention, if different than as described in the Examples. In general, the choice of vector depends on the size of the polynucleotide sequence and the host cell to be employed in the methods of this invention.

A "cassette" refers to a segment of DNA that can be inserted into a vector at specific restriction sites. The segment of DNA encodes a polypeptide of interest, and the cassette and restriction sites are designed to ensure insertion of the cassette in the proper reading frame for transcription and translation.

The terms "express" and "expression" mean allowing or causing the information in a gene or DNA sequence to become manifest, for example producing a protein by activating the cellular functions involved in transcription and translation of a corresponding gene or DNA sequence. A DNA sequence is expressed in or by a cell to form an "expression product" such as a protein. The expression product itself , e.g. the resulting protein, may also be said to be "expressed" by the cell. A polynucleotide or polypeptide is expressed recombinantly, for example, when it is expressed or produced in a foreign host cell under the control of a foreign or native promoter, or in a native host cell under the control of a foreign promoter.

A polynucleotide or polypeptide is "over-expressed" when it is expressed or produced in an amount or yield that is substantially higher than a given base-line yield, e.g. a yield that occurs in nature. For example, a polypeptide is over-expressed when the yield is substantially greater than the normal, average or base-line yield of the native

polypolypeptide in native host cells under given conditions, for example conditions suitable to the life cycle of the native host cells. Over-expression of a polypeptide can be achieved, for example, by altering any one or more of: (a) the growth or living conditions of the host cells; (b) the polynucleotide encoding the polypeptide to be over-expressed; (c) the promoter used to control expression of the polynucleotide; and (d) the host cells themselves. This is relative, and thus "over-expression" can also be used to compare or distinguish the expression level of one polypeptide to another, without regard for whether either polypeptide is a native polypeptide or is encoded by a native polynucleotide. Typically, over-expression means a yield that is significantly higher than a normal, average or given base-line yield. Likewise, a polypeptide is "under-expressed" when it is produced in an amount or yield that is significantly lower than the amount or yield of a parent polypeptide or under parent conditions. In this context, the expression level or yield refers to the amount or concentration of polynucleotide that is expressed, or polypeptide that is produced (*i.e.* expression product), whether or not in an active or functional form.

An expression product can be characterized as intracellular, extracellular or secreted. The term "intracellular" means something that is inside a cell. The term "extracellular" means something that is outside a cell. A substance is "secreted" by a cell if it delivered to the periplasm or outside the cell, from somewhere on or inside the cell.

"Isolation" or "purification" of a polynucleotide, gene, or protein refers to the derivation of the polynucleotide, gene or protein by removing it from its original environment (for example, from its natural environment if it is naturally occurring, or from the host cell if it is produced by recombinant DNA methods). Methods for polynucleotide, gene, or protein purification are well-known in the art, including, without limitation, electrophoresis, chromatography (including High Performance Liquid Chromatography or HPLC), and countercurrent distribution. For some purposes, it is preferable to produce the polynucleotide, gene, or protein in a recombinant system in which the polynucleotide, gene, or protein contains an additional sequence tag that facilitates purification. Alternatively, antibodies produced against the polynucleotide, gene, or protein or fragments derived therefrom, can be used as purification reagents. A purified polynucleotide or polypeptide may contain less than about 50%, preferably less

than about 75%, and most preferably less than about 90%, of the cellular components with which it was originally associated. A "substantially pure" enzyme indicates the highest degree of purity which can be achieved using conventional purification techniques known in the art.

A "control", "control value" or "reference value" in an assay is a value used to detect an alteration in, *e.g.*, transcriptional activity of a gene, the functional activity of an altered promoter, levels of a protein or mRNA detected in a sample taken from a patient or measured in a reconstituted system, or any other assays described herein. For instance, when studying modulation, *i.e.*, up- or down-regulation, of the transcriptional activity of an altered CD40L promoter sequence, the inhibitory/stimulatory effect of an agent can be evaluated by comparing the measured value of transcriptional activity to that of a control value. The control or reference value may be, *e.g.*, a predetermined reference value, or may be determined experimentally. For example, in such an assay, control or reference may be the transcriptional activity, *e.g.*, of the gene comprising the wild-type CD40L promoter; in the absence of the agent; in comparison with transcriptional activity with an agent having a known effect on altered CD40L promoter activity; or any other suitable control or reference. In a diagnostic assay, a reference or control value may be obtained by comparing *e.g.*, a nucleotide sequence, or a nucleotide or protein level measured, in a sample taken from a patient predisposed to or suspected of suffering from, a disease to a corresponding sequence or measured value of a sample taken from a healthy, or "control" individual.

An individual "at risk for", "predisposed to", or "susceptible to" a disease or condition means that the risk for the individual to contract or develop the disease or condition is higher than in the average population.

Abbreviations

Abbreviations used herein include:

Th (T-helper)

CD40L (CD40 Ligand)

sCD40L (Soluble CD40 Ligand)

RA (Rheumatoid Arthritis);

SLE (Systemic Lupus Erythematosus);
OA (Osteoarthritis);
JRA (Juvenile Rheumatoid Arthritis);
AVN (Avascular Necrosis);
ARMS (Amplification Refractory Mutation System);
EMSA (Electrophoretic Mobility Shift Assay);
ELISA (Enzyme-Linked Immunosorbent Assay);
FACS (Fluorescence Activated Cellular Sorting);
PBMC (Peripheral Blood Mononuclear Cells);
GAPDH (Glyceraldehyde-3-Phosphate Dehydrogenase);
ACR (American College of Rheumatology)
PMA (Phorbol-Myristate-Acetate)
CRE BP1 (Cyclic AMP Responsive Element Binding Protein 1)
NF-AT (Nuclear Factor of Activated T-cells)
PB (Peripheral Blood)
ST (Synovial Tissue)

Regulation of gene transcription

Proteins and enzymes can be made in a cell using instructions in DNA and RNA, according to the genetic code. "Transcription" is the process by which a DNA sequence or gene having instructions for a particular protein or enzyme is "transcribed" into a corresponding sequence of RNA. "Translation" is the process by which the RNA sequence is "translated" into the sequence of amino acids which form the protein or enzyme. Regulation of gene transcription involves regulatory elements in promoters and enhancers; structural or topological constraints placed on the regulatory elements, based on their location in the DNA double helix; the chemical state (e.g. methylation or acetylation) of the bases or DNA-associated molecules, such as histones; and the availability of the regulatory proteins and enzymes (transcription factors and polymerases) that initiate and mediate DNA transcription (63,91,97,98).

CD40 ligand

The gene encoding CD40L is composed of 5 exons and is located in the Xq26-27 chromosomal region (1,4,66). Several studies of the CD40L promoter have been published, with each supporting a requirement for binding of NF-AT to elements in the 5' promoter for induction of transcription (63,67,69). In addition, mice knocked-out for the NF-ATp transcription factor provide strong support for an essential role for that molecule in CD40L gene expression (99).

Cell surface CD40L protein. The discovery in several laboratories of the genetic basis of the hyper-IgM syndrome, characterized by impaired Ig class switching from IgM to mature Ig isotypes, emphasize the functional importance of CD40L-CD40 interaction for Th cell-mediated B cell differentiation (*see* above). Applicants have shown that SLE T cells have prolonged high level expression of CD40L after *in vitro* activation (28). Also, elevated CD40L levels may play a role in atherosclerosis and transplant rejection. Thus, systemic autoimmune diseases, characterized by production of high affinity IgG autoantibodies, are associated with increased or prolonged expression or activity of CD40L, resulting in persistent Th cell-mediated B cell activation and high local or serum levels of IgG. Moreover, lymphocytes from patients with clinically active SLE express some CD40L even in the absence of *in vitro* activation, and this CD40L is aberrantly expressed on CD8, as well as CD4, T cells (28). These results showed that the capacity for Th cell function is augmented in SLE and could be attributable to either multiclonal and persistent T cell activation by autoantigens, augmented T cell response to TCR-mediated signals, and/or impaired downregulation of lymphocyte activation. In addition to patients with SLE, patients with other systemic disorders, including those with RA, polyarteritis nodosa, hepatitis B or C, and other syndromes, were studied (28). Of those patients, there was variable expression of CD40L in response to stimulation of peripheral blood cells with PMA and ionomycin. Among those subjects who demonstrated prolonged expression of cell surface CD40L were several with RA or systemic vasculitis.

The effect on B cell activation of prolonged expression of CD40L on activated SLE Th cells was also studied. PBMCs from healthy subjects and SLE patients on costimulatory molecule expression on cocultured Ramos, CLL, or tonsil B cells, was studied (28). It was found that *in vitro*-activated SLE PBMCs induced significantly

higher CD80 expression on target B-cells than the low level of CD80 induced by activated normal PBMCs. Moreover, untreated SLE PBMCs induced higher levels of CD80 on the target B cells than did untreated normal cells, consistent with the higher levels of baseline CD40L expression in some SLE patients (28). These data show that the level of CD40L on circulating SLE lymphocytes, and that which persists on the SLE cell surface following activation, is functionally significant and contribute to excessive B cell activation.

Soluble CD40L protein. As the molecular structure of CD40L predicts a site suitable for enzymatic cleavage, it was predicted that sCD40L might also be increased in expression in systemic autoimmune diseases and that the soluble form might be functional (46). Two commercially available mAbs specific for different epitopes on the human CD40L molecule were used to establish an ELISA with specificity for sCD40L. Using this ELISA, it was shown that the mean level of soluble CD40L in sera from both clinically inactive and active SLE patients was significantly higher than the level in normal sera (46) and that the sCD40L in the active patients was considerably higher than in the inactive patients (46). Patients with other systemic diseases (with RA, anti-phospholipid syndrome, Lyme disease, or other non-SLE syndromes) demonstrated variable concentrations of serum sCD40L (*see* **FIG. 11**). Some had low or absent sCD40L by ELISA, but several patients with active RA, Wegener's granulomatosis, or polyarteritis nodosa showed increased levels. Of note, the mean level of sCD40L for 9 RA patients tested was 0.49 + 0.89 ng/ml compared to 0.025 + 0.04 ng/ml for healthy subjects. The specificity of the ELISA for sCD40L was confirmed in adsorption studies, in which the activity detected by the ELISA was removed by incubation of the sera with anti-CD40L mAb, but not by incubation with isotype matched anti-CD71 mAb. The presence of the 18 kD soluble form of CD40L in lupus sera was confirmed by western immunoblot (46). To test whether the soluble form of CD40L might be functional, it was first tested the capacity of recombinant trimeric sCD40L to induce B cell activation antigen expression on Ramos B cell line cells. Recombinant sCD40L, at 10 ng/ml, a level detected in some of the patient sera, induced increased CD95 expression on Ramos B cells. Similarly, some SLE sera increased B cell activation antigen expression on Ramos B cells, an effect that was inhibited by anti-CD40L mAb but not by control mAb (46).

These data document increased expression of the soluble form of CD40L in many patients with SLE and some with RA, show an association of sCD40L levels with disease activity, and raise the important issue of whether this soluble product can contribute to disease pathogenesis in vivo, by way of B cell, macrophage, dendritic cell, endothelial cell, or fibroblast/synoviocyte activation.

Thus, studies have documented increased and prolonged expression of cell surface CD40L and increased concentration of serum sCD40L in patients with systemic autoimmune disease, and have indicated that in SLE, the stability of CD40L mRNA may be prolonged (28, 46). In addition to modifications of mRNA stability, alterations in promoter sequence may affect transcription. As genomic DNA provides information related to promoter and intron sequences not available through study of cDNA, genomic DNA has been analyzed, including the CD40L immediate promoter sequences from healthy subjects and from individuals with systemic autoimmune disease (*see* Example 1).

Alterations in CD40L promoter

As sequence data derived from multiple genomic DNA samples have become available, variability in the regulatory regions of genes, most often representing genomic polymorphisms, has been described (105-108). Some of those sequence variations have been shown to confer altered transcriptional activity of the gene, with either increased or decreased production of mRNA and protein. Examples of promoter polymorphisms that alter gene regulation include the TNFα gene, with high or low producer variants correlating with particular promoter polymorphisms, and the type I collagen gene, in which a base change in a transcription factor binding site can decrease protein expression, contributing to clinical osteoporosis (105,108). Thus, altered promoter regions can be of major functional and clinical significance.

Applicants have surprisingly discovered alterations from the published sequence of the proximal promoter region of human CD40L in studies of genomic DNA isolated from ST and PB of patients with arthritis, SLE, or from healthy subjects (*see,* Example 1). First, the number of A's in the poly-A tract is variable among PCR-amplified genomic DNA subclones from all individuals studied, with the total length of this segment ranging

from 20-27 bp, even in a given individual. The length variability is localized to the 5' segment of the poly-A tract, with the number of A's ranging from 13 to 20. *See* **FIG. 3**. The mean length of the poly-A tract did not differ between healthy controls and arthritis patients. Poly-A tracts are subject to microsatellite instability, as discussed in the Background section, but since these tracts are most commonly found in introns or in the 3' untranslated regions of genes, the variability rarely has functional consequences. When poly-A tracts occur in a coding sequence, they may contribute to impaired or defective gene expression (78-80,84-87,90), and when localized to the 5' regulatory regions of genes, they may alter the regulation of transcription (91,92,94). Strategies for studying poly-A tract length variability and its functional consequences are provided below.

The second alteration from the published sequence of the CD40L proximal promoter noted in Example 1 was a substitution of a C for the A at position −125 from the transcription start site (*see* **FIG. 2** and **FIG. 5**). Position -125 from the transcription start site corresponds to residue No. 331 in **SEQ ID NO:1**, wherein residue No. 331 is an A, and in **SEQ ID NO: 2**, showing the A to C alteration at position No. 331. Alternatively, position -125 can be identified as 13 amino acids upstream from the CCTTT motif in SEQ ID NOS: 1 or 2. This alternative way of identifying position -125 is useful, for instance, when differences in poly-A length, deletions, insertions, or other mutations affect the numbering of the promoter residues. The A to C substitution results in 10 A's 5' of the substitution and 5 A's 3' of the substitution. In addition, an occasional subclone shows an extra C at various positions in the 6 A homonucleotide run, from positions −118 to −113 (*see,* **FIG. 3**). ABI Prism data were reviewed for each of the sequences shown in **FIG. 3**, and any sequences with "N's", suggesting unclear sequence data, in the −135 to −120 segment were excluded from analysis. In addition, PCR-amplified genomic DNA samples from some individuals were subcloned and sequenced at several time points, to exclude a role for technical artifacts of a particular sequence run in the poly-A tract alterations noted (see sequence dates noted in **FIG. 4**). Of greatest interest is the observation that of all samples studied, including genomic DNA samples from 23 healthy subjects, 7 SLE patients, 3 members of an extended Utah family, and from 46 patients with arthritis (predominantly those with RA), the A to C alteration has only been observed in samples from patients with arthritis and from 2 of the 3 Utah

family members. Also of interest are the results of sequencing genomic DNA from 4 male arthritis patients. As the gene for CD40L is located on the X chromosome, if the A to C alteration represents an allelic polymorphism, it would be expected that all subclones from a given male patient would have the same nucleotide at position -125. Indeed, ST64 shows 0/5 subclones with the A to C change; ST07 shows 4/4 subclones positive; and ST41 has 4/4 subclones positive (*see* **FIG. 4**). Taken together, the genomic DNA sequence data directed to the poly-A tract of the CD40L proximal promoter shows the existence of a genetic polymorphism defined by an A or a C at position -125. Most male patients show all subclones with the same nucleotide; 11 of 12 female patients with the A to C change have the altered sequence in some but not all subclones, suggesting that they are heterozygous for the alteration; and 2 of 3 members of a family have the altered sequence.

The length variability and alteration at position -125 of the poly-A tract in the proximal CD40L promoter represent features that present intriguing possibilities for altered conformation states, increased mutability of the surrounding nucleotides, and altered transcriptional regulation.

Described below are strategies to further study the A to C genetic polymorphism, to confirm the prevalence of the alteration among patient groups and healthy controls, and to define the proteins that differentially bind to the wild type and altered CD40L proximal promoter and nearby promoter regions.

Study of altered CD40L promoter in with RA and SLE patients

This section describes a strategy to study the altered nucleotide sequence in the proximal promoter of CD40L in an extended group of patients with RA, SLE, and healthy controls. The methods provided can also be used to identify an individual at risk for contracting or developing a disease in which elevated CD40L expression is a factor. Another strategy is provided in Example 1.

The study focuses on (1) the equivalent variability of the length of the poly-A tract in patients with RA, SLE and healthy control subjects; (2) the increased prevalence of an A to C alteration at position -125 in the poly-A tract of the CD40L proximal

promoter in patients with arthritis; and (3) whether the alteration of A to C represents an allelic polymorphism. The experimental approach is as follows:

Subjects. The study subjects for genomic DNA sequencing includes 50 healthy subjects, 50 patients with SLE, and at least 50 patients with RA. At least 50 PB samples are derived from patients with a clear diagnosis of RA, according to ACR criteria. In addition, PB samples from 50 patients with OA is studied to clarify whether the A to C alteration in the poly-A tract is increased in occurrence in patients with RA or OA or both. As several ST samples from patients with a diagnosis of OA, OA/RA, or AVN showed the A to C alteration (*see* above), it is possible that patients with OA who have significantly severe disease to warrant total joint replacement are genotypically different from OA patients with milder disease, who do not come to joint replacement surgery. Thus, OA PB samples include those from 25 patients who have had joint replacement surgery and 25 samples from patients with milder disease and no history or planned joint replacement surgery. Age, gender, and ethnic origin will be recorded for all study subjects. There is no exclusion of subjects on the basis of age, gender, or ethnic origin. In addition, 3 to 5 families of RA patients with the A to C alteration, and the extended Utah family from whom cell lines have been generated, will be studied. Also, to determine whether the expression of the altered A to C sequence differs between the peripheral blood compartment and the site of inflammation in the synovial membrane, 10 female patients will serve as donors of genomic DNA from both PB and ST. Thus, sufficient data is available to determine whether the A to C alteration is significantly different in occurrence in the study groups.

Thus, a sample may be taken from a patient, preferably a blood sample, and nucleic acid extracted from the sample. Nucleic acid can then be sequenced by any method know in the art. Non-limiting examples include:

Direct sequencing. Poly-A tract length and A to C nucleotide alteration is screened by direct sequencing. Direct sequencing of genomic DNA samples is performed across the 443 bp of the proximal CD40L promoter, either as an initial approach or after a preliminary screening. Based on the genomic sequence of CD40L, two primers, Pcd1 **[SEQ ID NO: 33]** and Pcd2 **[SEQ ID NO: 34]**, are synthesized (*See* **FIG. 2**). Genomic DNA is isolated and used as a template in PCR to amplify the 5' flanking sequence of

CD40L. The PCR product is subcloned into a T/A vector, positive clones picked, and plasmid DNA prepared and directly sequenced. At least 10 subclones are picked and sequenced for each sample studied. Once sufficient data is generated to determine whether length of the poly-A tract varies among study groups, the following screening approach is used to specifically detect the A to C sequence alteration.

Screening by ARMS method. ARMS is a two-step PCR amplification procedure (*see* **FIG. 12**). In the first PCR, a relatively long (443 bp) 5' flanking sequence of CD40L is amplified by primers Pcd1 and Pcd2. The PCR product is run on an agarose gel, the DNA band excised from the gel, passed through a spin column to remove the Pcd1 and Pcd2 primers, and then used as a template in the second nested PCR with primers Pcd3 **[SEQ ID NO: 35]** and Pcd4 **[SEQ ID NO: 36]**. Since Pcd4 is an altered sequence-specific primer, only the altered sequence (A to C) is amplified. The critical factor in screening altered sequences by the ARMS method is the annealing temperature in PCR. If it is set too low, non-mutated sequences can also be amplified, causing false positives. If it is set too high, then no product will be amplified. To optimize the annealing temperature, plasmid DNA samples with already known sequences can be used as positive and negative controls in PCR. The second PCR is performed as follows for each cycle: denaturing at 94°C for 1 minute; annealing at 60°C for 1 minute; and extension at 72°C for 1 ½ minutes. This amplification procedure is followed for 30 cycles. Positive results using ARMS screening, based on strong intensity bands, are confirmed by subcloning the sample's first PCR product, amplified by Pcd1 and Pcd2, and preparing and sequencing DNA.

Statistical analysis. Data comparing length of poly-A tract and occurrence of the A to C alteration at position −125 in patient and control groups is analyzed using the chi-square and Mann-Whitney tests. Through these experiments, it is determine whether the altered nucleotide sequences in the CD40L proximal promoter region, including the poly-A tract length variability and the A to C nucleotide change at position −125, represent a germline allelic variation, a result of insertions or deletions that occur in the context of DNA replication, or reflect other mutational events. It is also determined whether the altered sequence is enriched in the inflammatory milieu of the synovial membrane as compared to the peripheral blood, as might occur if cells expressing the A to C change

were preferentially expanded. Finally, a correlation between occurrence of the altered sequence and RA or destructive OA establishes that the A to C alteration is useful as a genetic marker for susceptibility to severe arthritis.

Binding of transcription factors

As CD40L is a critical molecule in T-B cell interaction, it should not be surprising that its expression is tightly regulated. Positive and negative cis-acting regulatory elements in gene promoters, including that of CD40L, bind transcription factors and contribute to control of gene expression. Several studies indicate that the level of CD40L mRNA expression parallels protein expression (47,64), with virtually no mRNA or protein expressed by the resting T cell. After TCR and CD28-mediated T cell activation, key transcriptional regulatory proteins move to the nucleus and induce CD40L promoter activity. Characterization of the proteins that bind to the CD40L promoter is mostly limited to one important study in the human system (63), and several murine studies, all concluding that NF-AT is essential for CD40L mRNA expression. A search in the established transcription factor binding site (TFSITES) data-base using the GCG program and the MatInspector version 2.2 program (available at World-Wide Web address transfac.gbf.de), permits identification of specific binding motifs (102). Review of these binding motifs suggested that proteins in addition to NF-AT are likely to bind to the 5' promoter. Of particular interest is that a TATA box is located just 5', and the consensus motif for the CRE binding protein is located just 3', to the poly-A tract. In addition, it was found that additional proteins can bind to an oligonuclotide that extends 22 bp 3' of the proximal NF-AT site.

It should be noted that in comparing mouse and human CD40L proximal promoter sequences, there is a high level of sequence conservation, with only 9 base differences, between mouse and human, if the poly-A tract is not considered (*see* **FIG. 8**). In addition to these base changes, the human sequence has lost 4 nucleotides when compared to mouse. This conservation is consistent with these promoter segments bearing important regulatory functions for CD40L transcription. In contrast to this high level of conservation, the poly-A tract in mouse and human bear considerable differences. While these promoter segments in the two species are clearly related, with the mouse

"poly-A tract" likely derived from an Alu element (103), 5 G's in the mouse sequence have mutated to A's in the human sequence, and a C has been gained in the human sequence. In addition, 11 nucleotides at the 3' end of the mouse sequence have been deleted in the human sequence. The change of A to C in the middle of the human poly-A tract confers a 4-6 fold increase in transcriptional activity (*see* **FIG. 7**). The location of the poly-A tract, spanning several potential transcription factor binding motifs, predicts that while the human poly-A tract may not itself bind essential transcription factors, variability in its length may modulate binding or function of factors binding nearby. Moreover, when an A is replaced with a C near the midpoint of the poly-A tract, a positive regulatory factor may be induced to bind or proteins that bind to adjacent motifs may do so more efficiently.

Characterization of transcription factors in patients with RA and SLE

This section describes a strategy to, using the same study subjects described above (*see* "Study of altered..."), characterize the transcription factors that bind to the altered promoter element as compared with wild type promoter sequence and nearby promoter elements. Positive and negative regulatory elements in the proximal promoter region of CD40L are identified, focusing on the approximately 150 bp 5' of the transcription start site, as well as their respective transcription factors. In addition, it is determined if the alteration of A to C at position −125 of the proximal promoter confers additional or altered binding of transcriptional regulatory proteins as compared to the wild-type sequence. Furthermore, to gain additional insight into the contributions of the poly-A segment to transcriptional regulation, the binding properties of oligonucleotides containing mouse or human poly-A segments are compared. The following experimental approach is used.

EMSA. DNA-protein binding complexes are determined by EMSA and supershift EMSA. EMSA will be used to identify specific binding sites in the CD40L promoter. A series of double-stranded oligonucleotide DNA probes, usually 25-30 bp, are synthesized to contain sequences of putative binding sites in the promoter region. A single strand oligonucleotide is synthesized by GIBCO-BR. Two reverse complementary single strand oligonucleotides are annealed and then radio-end-labeled with ^{32}P γ-ATP in the presence

of T4 polynucleotide kinase. Two to 3 mg of nuclear extracts isolated from peripheral blood T cells, unstimulated or stimulated with PMA and ionomycin for 2 hours, will be incubated with 1 ng end-labeled probe in the presence of poly dI:dC (*i.e.*, double-stranded polydeoxyiosine:polycytosine; Pharmacia) in a total volume of 20 µl at room temperature for 20 minutes. The reaction mixture is loaded and run on a 4.5% non-denaturing polyacrylamide agarose gel at 96 V for 40 minutes. The gel is dried and exposed to X-ray film at –80° overnight to demonstrate the DNA-protein binding complexes. To further confirm the specific binding, unlabeled probes at 50-fold molar excess are added to the reaction mixture to compete the binding with labeled probes.

From these data, it is determined if the substitution of a C for an A in the poly-A tract of the proximal promoter confers binding of a nuclear protein complex to a DNA probe spanning the poly-A tract. It is possible that the A to C change either modifies the binding capacity and transcriptional activity of neighboring elements, or the mutated poly A track may itself bind a functionally relevant transcription factor. Search of this segment, with the A replaced by the C, using the MatInspector version 2.2 program for identification of potential transcription factor binding motifs (available at World-Wide Web address http://transfac.gbf.de/) indicates that the alteration results in a potential binding site for proteins of the high mobility group (95). Such binding sites undergo significant bending to accommodate the binding protein, contributing to formation of a stable initiation complex (92). Several oligos are designed that either center the A to C change, or include 5' or 3' adjacent nucleotides to permit identification of nuclear proteins that bind to the poly-A tract segment, as well as the effect of the A to C change on binding of proteins to nearby sites. An oligonucleotide that substitutes the mouse poly-A segment for the human sequence is also used.

In addition, these data will extend currently available information regarding the binding motifs and associated proteins within the 150 bp 5' of the transcription start site. As noted herein, we have already determined that an oligonucleotide extending 22 bp 3' of the proximal NF-AT site binds a protein complex that persists in the presence of anti-NF-AT antibody. Recent supershift experiments, in which nuclear extracts from activated primary T cells are pre-incubated with specific antibodies prior to interaction with the labeled oligonucletides, suggest that this protein is a member of the Egr family (98).

To examine cell lineage specificity of proteins binding to the test probes, nuclear extracts are prepared from a panel of primary and cell line cells, including human Jurkat T cells, peripheral blood T cells, murine T cell lines, B cell lines (the Burkitt's lymphoma cell line Ramos and the CL-01 cell, representing a germinal center B cell), and non-lymphoid cell lines, such as Cos7 and HeLa cells. All cells are either cultured with medium alone for 2 hours, or with PMA and ionomycin, prior to isolation of nuclei.

When the segments of the CD40L proximal promoter that bind nuclear protein complexes are determined, a supershift assay is performed to identify the transcription factors which bind to the DNA sequences. Monoclonal antibodies (1-2 μg) specific to transcription factors, for example, anti-NF-AT, anti-fos, anti-jun, anti-ATF, or anti-CREB, will be added to the nuclear extracts for 2h at 4°C prior to adding the labeled DNA probes. If an antibody specifically binds an oligonucleotide-bound protein, after running the EMSA gel, the binding band is super-shifted to a higher position as the migration of the entire complex in electric field will be retarded. Antibodies to these and other transcription factors of interest are commercially available. If protein complexes bound to the poly-A tract or nearby nucleotides are not identified using the supershift approach, the bound complex is isolated and characterized. An increased or decreased activity of a motif to which an unidentified protein is bound can be studied in a luciferase assay (*see* below).

Mutational analysis. In the identified regions of the proximal promoter that bind nuclear extracts, *i.e.*, putative transcription factor binding sites, from activated T cells, mutations are introduced by site-directed mutagenesis using a PCR-overlapping method and confirmed by sequencing (104). EMSA assays is repeated with these mutated probes to determine the key nucleotides in protein binding.

Segments of the CD40L promoter affecting activity

This section describes a strategy to identify and study the segments which affect CD40L promoter activity and fragments of the CD40L promoter that are functional, i.e. allow or promote transcription. Based on the genomic DNA sequence of CD40L, and the information on nuclear protein binding motifs generated in the previous section, primers are designed and the 5' flanking sequence of CD40L amplified in order to

generate fragments of different length. For example, the entire 1.5 Kb fragment, a fragment 443 bp 5' upstream of the transcription initiation site, and a series of 5' deletion segments are produced. These fragments are subcloned into the pGL-2/basic luciferase reporter vector and transiently transfected into human Jurkat T cells or into ConA-activated primary peripheral blood T cells, as recently reported (109). By comparing relative light units, the segments that are the main contributors to the promoter activity, and regions which can enhance or inhibit the promoter activity, are localized. *See* **FIG. 10.** In particular, active promoter fragments containing a poly-A tract with either A or C at position -125 are studied and compared. In addition, protein binding motifs and mutated variants are investigated, and luciferase assays in which the human poly-A tract is substituted by the mouse poly-A segment and transfected into either human or mouse activated T cells or T cell line cells, are performed. *See* also Example 4.

Furthermore, negative regulatory elements are studied. There is minimal constitutive production of CD40L mRNA or protein under baseline cellular conditions, suggesting that its promoter may be under the influence of negative regulatory factors in the absence of stimulation. In contrast to other T cell genes, such as CD25, which remain turned on for days after induction, CD40L mRNA is only briefly expressed. The MatInspector version 2.2 program indicates that the promoter sequence just 3' of the poly-A tract contains a possible binding motif for the repressor protein E4BP4, as well as the CRE-bp consensus sequence (110-113). A negative regulatory protein identified is overexpressed in Jurkat cells to assess its repressive function in the luciferase assay.

After activation of CD4+ T cells with anti-CD3 mAb, (or other T cell mitogens, such as PMA, ionomycin and Con A), CD40L expression peaks at 2-6 hours and has nearly returned to the basal level after 12 hours (28). This tight regulation may reflect the requirements for activation of positive regulatory factors, as well as a possible autoregulatory mechanism that represses transcription within several hours after induction. CD40L promoter activity response to these stimuli, and whether the response occurs in a time-dependent fashion, is studied. After transient transfection, transfectant Jurkat T cells is stimulated with PMA, ionomycin, and Con A at different time points, such as 1, 2, 6, and 12 hours. The cells are then lysed, and the transcription activity in the lysates is measured by a luciferase assay. These kinetics experiments are performed using

constructs expressing the wild type and the altered CD40L promoter poly-A track.

Comparative CD40L transcription

This section describes a strategy to compare CD40L transcription in cells from healthy subjects with that of RA patients with wild-type promoter poly-A sequence and RA patients with homozygous and heterozygous alteration of the poly-A sequence.

RA patients shown to express the altered CD40L promoter (C/C or A/C) sequence are compared to patients with the wild-type (A/A) sequence and with healthy subjects with the wild-type sequence for CD40L transcriptional activity. Competitive PCR is used to measure CD40L mRNA in unstimulated PBMC and in cells activated with PMA and ionomycin, ConA, or anti-CD3 monoclonal antibody. Total cellular RNA is extracted from the samples by acid guanidinum thiocyanate-phenol-chloroform extraction and reversed transcribed into cDNA using the SuperScript First Strand cDNA synthesis kit (GIBCO-BRL Life Technology Inc., Gaithersburg, MD). Relative quantities of CD40L mRNA is determined by competitive mimic RT-PCR. (*see* Examples). This analysis shows whether those patients who demonstrate the A to C alteration in the CD40L promoter can generate higher levels of CD40L mRNA after activation of their PBMC cells in vitro.

Cell populations displaying altered CD40L expression.

CD40L is predominantly expressed on activated CD4+ T helper cells. However, some studies, including our own, indicate that CD40L can be found on CD8+ T cells, and others have found the molecule on B cells, activated platelets, and other cell populations. In view of the important role of CD40L in promoting B cell expansion, increased expression of co-stimulatory molecules, and subsequent T cell activation, increased CD40L expression by a given cell might contribute to preferential expansion of that cell. If the alteration of A to C in the poly-A tract of the proximal promoter confers increased transcriptional activity, that altered promoter sequence might confer preferential expansion of the T cells with the C genotype. It is also possible that non-T cell populations with the C genotype might preferentially express surface CD40L.

This section describes a strategy for determination of cell populations which contain the A to C mutation at position -125, (corresponding to residue 331 of **SEQ ID NO: 2**), of the CD40L proximal promoter **[SEQ ID NO: 1]**. As CD40L is encoded on the X chromosome, cells from a female heterozygous for the proposed genetic polymorphism will either express the A or C poly-A tract sequence in a given cell. At the population level, any advantage that C-expressing cells have compared to A-expressing cells should be discernable if the advantage is significant.

PB and SF samples from several such patients is fractionated into CD4 or CD8 T cells, CD69+ and CD69- subsets, CD19+ B cells, and CD14+ monocytes. In addition, ST samples is digested with collagenase, hyaluronidase and DNAse, mononuclear cells isolated and similarly fractionated, and the remaining material cultured for 7 days to obtain fibroblastoid synovial cells. In addition, ST fragments are cultured with IL-2 or IL-15 for 7 days to derive T cell populations whose growth is promoted in the context of the ST matrix. All cell populations are used for preparation of genomic DNA, PCR amplified using the ARMS screening method, and relative expression of the A to C change in the poly-A tract determined. A skewing toward expression of the promoter sequence with a C at position -125 would suggest that those cells have a survival or proliferation advantage. Such a result should be followed-up with appropriate functional analysis, depending on the cell populations that give the skewed results.

Correlation between CD40L transcription and surface expression

Several studies suggest that CD40L transcription correlates with CD40L cell surface expression. Thus, the altered promoter sequence would affect the level of cell surface CD40L inducible in vitro and sCD40L expressed in vivo. Differences among individuals in production of CD40L is most readily discerned by measuring the soluble form in serum or plasma. While sCD40L is hardly detectable in sera from healthy subjects, levels are highly significantly increased in patients with SLE, as well as those with RA and other systemic vasculitis syndromes. Individuals with the A to C change at position -125 of the CD40L promoter would thus express higher levels of CD40L cell surface protein and sCD40L in serum. The following section describes a strategy to study

the relationship between cell surface levels of CD40L inducible *in vitro*, and the levels of sCD40L *in vivo*.

Cell surface CD40L. Cell surface CD40L is measured on unstimulated or PMA and ionomycin-stimulated CD4+ T cells from healthy subjects and from RA patients with the A/A, A/C, or C/C genotypes, as inferred from sequencing of at least 10 subclones from genomic DNA amplified with the Pcd1 and Pcd2 primers. Cells are stained and analyzed by two-color immunofluoresence for CD40L and CD4 at 6 and 36 hours after initiation of culture.

Soluble CD40L expression. Sera from 25 healthy subjects and sera and SF from 50 RA patients with the A/A, A/C, or C/C genotypes, is assayed for sCD40L by ELISA. Since many of the RA fluids contain rheumatoid factor which has the potential to react with the antibodies used in the ELISA and produce an falsely high result, all fluids are depleted of Ig prior to assay by passing over a Staphylococcus protein A column. Briefly, microtiter plates are coated overnight with 100 ng/well mouse anti-CD40L mAb (TRAP1 clone, Pharmingen, San Diego, CA) and blocked with 1% Carnation milk in PBS-Tween. Fifty ml of either serum or SF samples, diluted 1:50 or 1:100 in PBS, or a range of concentrations of recombinant trimeric human CD40L, are added to the microwells in triplicate. After overnight incubation and washing, the assay is developed with alkaline phosphatase-labeled anti-CD40L mAb (Ancell, Bayport, MN), reactive with a different epitope of CD40L than the coating mAb. Relative concentration of soluble CD40L in each sample is determined after developing the reaction with substrate, and comparing sample O.D. reading to the standard curve.

Role of homonucleotide runs on promoter function

Variability in the length of the poly-A tract in the proximal CD40L promoter, in a region rich in potential transcription factor binding sites, raises questions regarding the basis of the length differences, as well as the effects of that variability on promoter function.

Consideration of the mouse and human CD40L promoter sequences reveals that in both species the proximal promoter is marked by an interruption of 5' and 3' regions by an adenine-rich segment (Figure 8). Such poly-A tracts are common in the 3'

untranslated regions of genes, are most likely derived from retrotransposition of an Alu element, and are thought to play an important role in regulating the stability and persistence in the cytoplasm of mRNAs (103). In contrast, poly-A tracts are rare in 5' promoters, and there are few rules that can be gleaned regarding their possible functions in that context. In the mouse CD40L gene, this region is characterized by 20 A's with interpersed G's, and 10 bases are situated 3' to the A rich region. In man, this 10 nucleotide segment is no longer seen, and the A-rich segment has lost all guanines, with only a C breaking a string of 22 A's in the published sequence. The A rich segment in mouse and man are clearly related, but significant alterations have occurred from one to the other species. In contrast, the promoter regions both 5' and 3' to the poly-A tract are highly conserved in both mouse and human and are rich in potential transcription factor binding sites. While mouse and human CD40L promoters have not been studied side by side, nor has the magnitude or kinetics of CD40L expression been directly compared in the two species, it is likely that the variable features of the A-rich tract alter the regulation of gene expression.

Among human genomic DNA samples, including multiple DNA subclones from a single sample, considerable variability in the length of the poly-A tract in the proximal CD40L promoter is observed (see Table 1). In both healthy subjects and patients, the length of the poly-A tract, from the universally conserved 5' ATT to 3' CCTTT, varies from 20 to 27 bp. No apparent relationship between tissue source of genomic DNA (synovial tissue or peripheral blood) or diagnosis of patient and poly-tract length is perceived. The basis for this variability in segment length is not clear. While genomic DNA is being studied and the first assumption that must be made is that the sequences obtained are encoded in the germline, it is difficult to imagine, although possible, that multiple replicate copies of the CD40L gene, each with a different length of poly-A tract, are thus encoded.

Alternatively, it is possible that with DNA replication, poly-A sequences are vulnerable to additions or deletions that might contribute to variable segment length in progeny cells. Abundant precedent is available to support the prediction that the poly-A tract in the CD40L proximal promoter is unstable and subject to deletions or additions based on frameshift or other errors during DNA replication. Eukaryotic genomes contain

many regions of DNA in which either single, double, triple, or greater numbers of nucleotides are repeated in tandem. These stretches of repeated bases are termed micro- or minisatellites, depending on their length. Micro and minisatellites are highly unstable and vulnerable to being replicated with impaired fidelity during DNA synthesis (89). Although these events have been best documented in the setting of deficiencies in mismatch repair genes, that are ordinarily responsible for correcting the mismatches that occur when the two DNA strands do not anneal properly, that repair system may not be involved in maintaining stability of homonucleotide repeats of 16-20 bp in length (88,89,114). Nucleotide repeats within the coding sequence of genes are vulnerable to generating functionally significant changes in sequence in that setting, as has been observed in the factor IX and TGFb receptor type II genes and the APC gene in certain malignancies (78,81-87,90). Of particular interest, poly-A tracts, as in the TGFb receptor gene, are particularly vulnerable, and cannot only themselves undergo changes in length, but can serve as a hypermutable site for neighboring nucleotides (85,115). Several examples of poly-A tracts in gene promoters suggest that their variability may also affect promoter function, as discussed in the Background section. Whether the origin of poly-A tract variability is genetic or a result of somatic alterations will be investigated by searching for multiple genomic copies of the CD40L gene and by analyzing poly-A tract length in clonal T cell populations.

The following rational is the basis for characterizing the activity of CD40L promoters containing poly-A tracts of varying length. Variability in the number of A's in the poly-A tract may alter the efficiency of binding of transcription factors to neighboring binding sites and may alter the efficiency of transcription. Proteins binding to both 5' and 3' sides of the poly-A tract are likely to need to appropriately associate to trigger transcription initiation and progression. When these motifs are brought closer together or stretched farther apart by the intervening A's, the topology of DNA may be changed. For example, proteins that should be binding in tandem on the double helix may be placed on opposite sides of the double helix.

Whether the origin of poly-A tract variability is genetic or somatically derived can be investigated by searching for multiple genomic copies of the CD40L gene and by analyzing poly-A tract length in clonal T cell populations. The transcription factors

bound and the activity of CD40L promoters containing poly-A tracts of varying length can also be studied. The following section describes a strategy for characterizing the basis of the variable length of the poly-A tract in the CD40L proximal promoter. The poly-A tract length variability in a clonal population, the effect of DNA replication and cell division on poly-A tract length, and transcription factor binding to variable length poly-A tracts, are analyzed using the following experimental approaches.

Poly-A tract length variability in a clonal population. The Jurkat cell line is used as a clonal cell population that should express a poly-A tract of uniform length if insertions or deletions during DNA synthesis play no role in determining the CD40L promoter sequence. As the Jurkat cell line is derived from a male (ATCC catalogue), it should have only one functional copy of CD40L in its genome. Genomic DNA is isolated from Jurkat, PCR amplified using the Pcd1 and Pcd2 primer set, and PCR products subcloned and sequenced. At least 20 subclones are directly sequenced to yield data for the 443 bp proximal promoter. Only one poly-A tract length detected among the 20 sequences obtained would indicate that insertions or deletions do not modify poly-A tract length during cell replication. Such an observation would be confirmed using other clonal cell lines. Should more than one poly-A tract length be identified among the subclones, two explanations would be that (1) more than one CD40L gene is expressed in each genome, or (2) DNA replication and cell division result in alterations in poly-A tract length, of which the latter one is the most plausible.

The possibility that the CD40L gene is reduplicated in the germline, with several tandem or widely distributed copies each expressing a poly-A tract of different length, can be explored as follows. Primer sets are designed such that the 5' primer amplifies the proximal promoter, just 5' of the poly-A tract, and the 3' primer amplifies the 5' end of the second intron (3' of exon 1). This approach relies on the assumption that if the CD40L gene is replicated in multiple copies in the genome, it is likely that intron sequences will not be identical among the various gene copies. Thus, the primer set that spans the poly-A tract, exon 1, and part of intron 2 would be predicted to amplify products of restricted poly-A length, while primer sets amplifying only the proximal promoter will generate sequences of various poly-A lengths. Should variable poly-A lengths continue to be generated, even when using a spectrum of primer sets that amplify

at different segments past the first exon, it is more likely than not that the variable poly-A tract lengths derive from somatic variablility.

Effect of DNA replication and cell division on poly-A tract length. To address the possibility that DNA replication and cell division can generate variable lengths of poly-A tract in the CD40L promoter, both Jurkat cell line cells and peripheral blood T cells from male donors are used. In using male donors, the source of DNA containing the CD40L promoter is limited to one X chromosome. Jurkat and primary T-cells are seeded in microwells at low concentration to generate cultures with relatively limited T-cell heterogeneity. Primary T-cell cultures are supported with PHA and IL-2 to expand the initial seeded cells. Genomic DNA is isolated from aliquots of cells harvested at 2, 4, 6, 8, and 10 days after initiation of culture, and CD40L proximal promoter PCR amplified, subcloned and sequenced. At least 10 subclones are sequenced for each time point. If DNA replication contributes to insertions and deletions that result in poly-A tracts of variable length, the degree of variability of poly-A tract length among sublones sequenced would increase with each time point studied.

Activity of CD40L promoters containing varying length poly-A tracts. First, the effect of variable length of poly-A tract on transcription factor binding is studied. Mutant double stranded oligonucleotide constructs are made that span the poly-A tract of the 5' proximal promoter, with 4, 8, 12, 14, 16, 20, or 24 A's replacing the 16 5' A's of the wild type polyA tract and with the oligonucleotide including the 5' and 3' putative transcription factor binding motifs. These oligonucleotides are ^{32}P-labeled and used in EMSA studies, as described above. If variable binding of nuclear extracts from activated T cells to probes containing different numbers of A's is detected, the strongest and weakest binding oligonucleotide are selected for further study. Supershift assays and semi-quantitative studies of dilutions of nuclear extracts are used to determine if differences in binding are qualitative or quantitative. This will indicate whether poly-A tracts of varying length bind different proteins, or whether they bind the same proteins with different efficiencies.

Next, the effect of variable length of poly-A tract on promoter activity is studied. Mutant double stranded oligonucleotide constructs are made that span the 5' proximal 443 bp of the CD40L promoter, with 4, 8, 12, 14, 16, 20, or 24 A's replacing the 17 5'

A's of the wild type poly-A tract. These constructs are transiently transfected into Jurkat cells or activated primary T cells and luciferase activity measured after 48 hours of culture with medium or PMA and ionomycin added during the last hour of culture. It is predicted that a number of A's both less and greater than the typical 16 A polynucleotide tract (between the 5' ATT and the C at position -119) will decrease the transcriptional activity of the promoter. These results and the parallel EMSA data are used to make predictions regarding the role of poly-A tract length on efficiency of binding of nuclear proteins and transcriptional efficiency of the proximal promoter.

Antibodies to altered CD40L promoter

According to the invention, altered CD40L proximal promoter polypeptides produced recombinantly or by chemical synthesis, and fragments or other derivatives or analogs thereof, including fusion proteins, may be used as an immunogen to generate antibodies that recognize the altered CD40L proximal promoter. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments, and an Fab expression library. Such an antibody is preferably specific for altered CD40L promoter from mammals, including but not limited to, humans.

Various procedures known in the art may be used for the production of polyclonal antibodies to the altered CD40L promoter or derivative or analog thereof. For the production of antibody, various host animals can be immunized by injection with the altered CD40L promoter, or a derivative (*e.g.*, fragment or fusion protein) thereof, including but not limited to rabbits, mice, rats, sheep, goats, etc. In one embodiment, the altered CD40L promoter or a fragment thereof can be conjugated to an immunogenic carrier, e.g., bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH). Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (*Bacille Calmette-Guerin*) and *Corynebacterium parvum*. Antisera may be collected at a chosen time point after immunization, and purified as desired.

For preparation of monoclonal antibodies directed toward the altered CD40L promoter, or fragment, analog, or derivative thereof, any technique that provides for the production of antibody molecules by continuous cell lines in culture may be used. These include but are not limited to the hybridoma technique (127), the trioma technique, the human B-cell hybridoma technique (128, 130), and the EBV-hybridoma technique to produce human monoclonal antibodies (129).

The foregoing antibodies can be used in methods known in the art relating to the localization and activity of the PAMP polypeptide, *e.g.*, for Western blotting, imaging altered CD40L promoter *in situ*, measuring levels thereof in appropriate physiological samples, etc., using any of the detection techniques mentioned above or known in the art. Such antibodies can be used to identify proteins that interact with the altered CD40L promoter, and to detect conformational or structural changes in the altered CD40L promoter. In a specific embodiment, antibodies that agonize or antagonize the activity of altered the CD40L promoter polypeptide can be generated.

Assay for evaluating inhibition and/or stimulation of altered CD40L promoter function

Identification and isolation of the altered CD40L promoter provides for development of screening assays, particularly for high throughput screening of molecules that up- or down-regulate, *i.e.,* inhibit or stimulate, the translation activity of the altered CD40L promoter, *e.g.*, by permitting expression of CD40L in quantities greater than can be isolated from natural sources, or in indicator cells that are specially engineered to indicate the amount or activity of CD40L expressed via an altered promoter sequence after transfection or transformation of the cells, or by inhibiting the transcription of CD40L by interacting with the altered promoter sequence. The present invention contemplates screens for small molecule ligands or ligand analogs and mimics, as well as screens for natural ligands that bind to and up- or down-regulate the translational activity of the altered CD40L promoter *in vitro* or *in vivo*.

Any screening technique known in the art can be used to screen for compounds which up- or down-regulates the translation activity of the altered CD40L promoter. For instance, a screening assay can be based on measurement of the amount or formation rate

of transcribed CD40L mRNA by a suitable method, the luciferase assay described above, or transcription of the CD40L gene from an altered promoter resulting in the formation or release of a reporter molecule which can be easily measured. Generally, a screening assay involves contacting the altered promoter sequence with a compound which interacts or otherwise affect the promoter function and/or conformation. Preferably, the altered CD40L promoter sequence is linked to cDNA encoding for a reporter protein, or CD40L or a fragment thereof, or another polypeptide or protein. The transcriptional activity of the altered promoter is measured in the presence of the compound, and compared to a control value. This control value could be, for example, transcriptional activity of the altered promoter in the absence of the compound, transcriptional activity of the wild-type CD40L promoter in the presence of the compound, transcriptional activity of the altered promoter in the presence of a compound with a known effect on transcriptional activity, or another theoretically or experimentally derived value.

CD40L diagnostic assay

The present invention provides for a novel method to diagnose and/or confirm autoimmune diseases, especially RA, by detecting an alteration of the CD40L promoter sequence. For instance, in one embodiment, a blood sample or tissue sample, preferably a blood sample, is taken from the patient diagnosed with, predisposed to having, or suspected of having, RA or another disorder in which elevated CD40L is a factor, and nucleic acid is extracted from the sample and sequenced (*see* below). Preferably, the sequence is then compared to suitable control sequences, such as, *e.g.*, **SEQ ID NO: 1**, while compensating for any differences in poly-A length, to see whether there is an A to C substitution at position -125 (corresponding to residue 331 of **SEQ ID NO: 2)** of the CD40L proximal promoter **[SEQ ID NO: 1]**.

In another embodiment, a blood or tissue sample which contains cells is taken from an individual at risk for or predisposed to having RA or another disorder in which elevated levels of CD40 is a factor. The nucleic acid can be extracted, and/or a level of transcriptional activity of the CD40L promoter measured (*see* Example 2). Preferably, the measured value of transcriptional activity is compared to a control value to evaluate whether there is a substantial difference, in which case the individual is at risk for the

disease being evaluated. The control value can be, for instance, the transcriptional activity of a CD40L promoter in a sample taken from a healthy control individual; the average CD40L transcriptional activity measured in a population of healthy individuals; or another suitable control value. Preferably, a similar type of sample is taken from the control individual as the individual at risk for a disease, and the two samples processed using substantially similar procedures. In one embodiment the sample and the control sample are analyzed in parallel to minimize the influence of variations in experimental conditions. In another embodiment, a control sample is analyzed prior or subsequent to the sample taken from the individual at risk for the disease being investigated.

In an alternative embodiment, CD40L mRNA or CD40L is isolated from an individual to investigate, for example, whether CD40L mRNA transcription or CD40L expression levels differ from typical levels, *i.e.*, control levels measured in healthy individuals, in procedures similar to those described above.

Poly-A tract length and A to C nucleotide alteration can be screened by direct sequencing, ARMS, or any other sequencing method known in the art. Sequencing of genomic DNA samples can be performed across the 443 bp of the proximal CD40L promoter, either as an initial approach or after a preliminary screening. Based on the genomic sequence of CD40L, two primers, Pcd1 **[SEQ ID NO: 33]** and Pcd2 **[SEQ ID NO: 34]**, can be synthesized (*See* **FIG. 2**). Genomic DNA can be isolated and used as a template in PCR to amplify the 5' flanking sequence of CD40L. The PCR product can then be subcloned into a T/A vector, positive clones picked, and plasmid DNA prepared and directly sequenced. In one embodiment, at least 10 subclones are picked and sequenced for each sample studied.

The knowledge derived from the procedures described above would allow for better diagnostic procedures for identifying individuals at risk for, susceptible to, or predisposed to RA or other diseases in which elevated CD40L transcription/expression is a factor, and the role of the proximal promoter elements of the CD40L gene in its transcriptional regulation. The correlation between the altered promoter sequence and RA, identification of the cis-acting regulatory elements in the wild type and altered CD40L promoters and their specific transcription factors, and information about how alterations in the proximal promoter modulate CD40L gene expression, will provide for

a better understanding of the causes and progression of RA and other autoimmune or inflammatory diseases, as well as other CD40L-related diseases or conditions, as well as novel therapeutic strategies for treating such diseases or conditions.

EXAMPLES

The present invention will be further understood by reference to the following examples, which are provided as exemplary of the invention and not by way of limitation.

Useful techniques for these Examples include:

ARMS. As sequence variability in the poly-A tract in the proximal promoter does not confer any restriction enzyme site change, it is not analyzed by the restriction fragment length polymorphism (RFLP) method. An alternative approach to screening for alteration in promoter sequence is by the Amplification Refractory Mutation System, or ARMS analysis (100). This is a two-step PCR amplification procedure, as shown in **FIG. 12**. The ARMS method has been used successfully in screening for human TCR Vβ17 allelic variations in previous studies (101).

EMSA. The electrophoretic mobility shift assay (EMSA) is a tool to identify DNA-nuclear protein complexes (63).

The Luciferase assay (63), the DEAE-dextran electroporation method (63), the ABI Prism technique, and other techniques used herein, are known in the art

EXAMPLE 1

Analysis of the CD40L proximal promoter in patients with systemic autoimmunity

In this example, genomic DNA including the CD40L immediate promoter sequences, from healthy subjects and from individuals with systemic autoimmune disease, was analyzed. Peripheral blood samples from 23 healthy subjects, 7 SLE patients, 11 RA patients, and 3 samples from an extended Utah family (see World-Wide Web address at locus.umdnj.edu/nigms, family No. 1331 of the CEPH/Utah pedigree sets repository, No. GM06983), were used for isolation of genomic DNA, followed by DNA sequencing of the 443 bp proximal CD40L promoter (126). In addition, synovial

tissue samples from 32 patients with RA, 2 patients with juvenile arthritis, 2 patients with an assigned diagnosis of OA/RA, 1 patient with avascular necrosis (AVN), and 1 patient with osteoarthritis (OA) were similarly analyzed. (*See* TABLE 1 for gender and ethnicity information).

Two primers, Pcd1 and Pcd2 were synthesized, and genomic DNA used as a template in PCR to amplify the 5' flanking sequence of CD40L (from Genbank Accession L47983). The PCR product was subcloned into a T/A vector, positive clones picked, and plasmid DNA prepared and directly sequenced. **FIG. 2** shows the 5' flanking sequence alignment for wild-type and altered CD40L. The promoter regions amplified by Pcd1 and Pcd2 are indicated by an underline, along with those amplified by a second primer set, Pcd3 and Pcd4, indicated by a double underline (*see* **Example 3**). Position -125, altered from an A to a C in some samples, is indicated by an *.

These data identified alterations in the proximal CD40L promoter from some DNA samples that may have the potential to modify CD40L promoter function. All altered sequences observed are localized in a poly-A tract located at -135 to -113 5' of the transcription start site. The poly-A site comprises 16 A's (-135 to -120), a C at position -119, and 6 A's (-118 to -113) (Genbank Accession No. L47983). The first class of alterations noted is characterized by variability in the length of this poly-A tract. All samples studied, including those from healthy subjects and patients with RA or SLE, show variable length of the poly-A tract in multiple subclones sequenced, with most of the variability localized to the 5' poly-A segment (representative sequences shown in **FIG. 3**). In contrast to the published 16 A's at -135 to -120, our data documented a range of 13-20 A's, resulting in a length of the total poly-A tract (-135 to -113 segment) that varies from 20-27 bp among all subclones sequenced. The mean poly-A tract length for arthritis patient samples (23.3 ± 0.68) does not differ from the length in normal subjects (23.3 ± 0.59). There was no apparent difference in the degree of poly-A tract length variability between ST and peripheral blood samples from the arthritis patients.

The second and more intriguing alteration in the proximal promoter was characterized by a nucleotide substitution of A to C at position -125 inside the poly-A. The results of direct sequencing of the CD40L proximal promoter from 5 arthritis ST and 2 control peripheral blood samples are shown in **FIG. 3** (representative poly-A tract

sequences). All samples demonstrated the consensus ATT 5' of the poly-A tract and CCTTT 3' of the poly-A tract. Variability in the length of the poly-A tract was observed in all individuals studied, and the substitution of a C for an A at position -125 was detected in genomic DNA samples isolated from 9 of 38 ST samples and 6 of 11 PB samples from patients with arthritis (*See* **FIG. 4**). Of these, 2 patients were donors of both synovial and blood samples, with both tissue and blood giving concordant results. A third ST/PB pair (from patient 41) gave discordant results and sequencing on the PB is being repeated. The patients with the A to C substitution included 9 with a diagnosis of RA, 1 with JRA, 1 with OA/RA, 1 with OA, and 1 with AVN. In some samples, the genomic DNA was sequenced across the proximal promoter on both strands, with results confirming the alteration ("T" to "G" on the opposite strand). Shown in **FIG. 5** are representative ABI Prism data demonstrating wild-type and altered poly-A tract sequence in 2 subclones from a synovial tissue sample taken from an RA female. Genomic DNA was amplified by PCR using Pcd1 and Pcd2, subcloned, and sequenced. The bottom panel of **FIG. 5** demonstrates a poly-A tract expressing the altered A to C at position -125.

Shown in **FIG. 13** is the results from screening for the A to C alteration in the CD40L proximal promoter by the ARMS method. Eight synovial tissue samples from arthritis patients were screened with ARMS and compared with a positive control sample (with A to C substitution). Alteration of A to C at position -125 of the poly-A tract was confirmed by direct sequencing of two samples (ST31 and ST30).

In contrast to the CD40L proximal promoter sequences derived from arthritis patient samples, no alterations of A to C at position -125 were noted in PB samples from 23 healthy subjects or 7 SLE patients. The sequence of genomic DNA from 3 lymphoblastoid cell lines generated from members of an extended Utah family was studied, and 2 of 3 family members studied demonstrating the A to C alteration in the poly-A tract. While the health status of the donors of these cell lines is not known, the data suggest that more extensive family studies may support the designation of the A to C variation as an allelic polymorphism.(See **FIG. 6**).

TABLE 1

Ethnicity and Gender of Human Subjects Studied

Subjects:	Caucasian	African American	Hispanic	Asian	Other (or not known)	Total
Female	35	3	8	8	18	72
Male	7	0	3	2	0	12
Total	42	3	11	10	18	84

EXAMPLE 2

Time course of induction and expression of CD40L mRNA

In this Example, the time course of induction and expression of CD40L mRNA was studied to investigate whether the prolonged cell surface CD40L expression observed on T cells from patients with systemic autoimmune diseases, as well as the increased circulating levels of sCD40L, would be associated with increased or prolonged cellular expression of CD40L mRNA in those patients.

Northern blot and competitive mimic polymerase chain reaction (PCR) assays were established for human CD40L in order to assess the time course of induction and expression of CD40L mRNA after activation of PBMC with PMA and ionomycin. In the case of the northern assays, cellular RNA was assayed using a ^{32}P-labeled CD40L probe, in parallel with a probe specific for the stable and abundant cellular mRNA for GAPDH. In the case of the competitive PCR, cDNA was reverse transcribed from cellular RNA, and a range of concentrations of a molecular construct that contained a nucleotide sequence derived from CD40L was included in each test PCR reaction for CD40L cDNA.

The results of one such study is shown in **FIG. 9**. PBMC from a healthy subject (left side of gel) or a patient with SLE (right side of gel) were incubated for one hour with PMA and ionomycin. Replicate cultures were then either cultured for an additional hour without any further additions (top panel), or with actinomycin D (5 μg/ml; bottom panel). RNA was prepared from cell extracts and reverse transcribed into cDNA, and PCR reactions were performed in the presence of a range of concentrations of a mimic construct (from residue base 418-1271 from Genbank Accession No. L07414; CD40L mRNA), containing a portion of the CD40L DNA sequence. For each set of PCR reactions shown, the lower band indicates the product of the amplified mimic construct and the upper band indicates the product of the test cDNA. The concentration at which

the density of the upper test band exceeds the density of the lower mimic band indicates the concentration of mimic that cannot out-compete the test cDNA for amplification by the CD40L PCR primers. In this experiment, in the absence of actinomycin D, the control cDNA out-competes the mimic at a mimic concentration of 10^{-3} attomoles/ml, while the SLE cDNA out-competes the mimic at a mimic concentration of 10^{-2} attomoles/ml. Therefore, the SLE cDNA contains roughly 10 times more CD40L cDNA than does the control cDNA. For the lower panel, representing cells cultured with actinomycin D added at the 1 hour time point, no CD40L cDNA is detected for the control, while a mimic concentration between 10^{-2} to 10^{-3} out-competes the SLE cDNA.

When comparing CD40L mRNA expression in PBMC isolated from SLE patients or normal subjects, data from both northern blot and competitive PCR assays showed that while the maximum CD40L mRNA expression was observed at 1-2 hours in both subject groups, the relative expression of CD40L mRNA, compared to GAPDH mRNA, was greater in SLE than control subjects. In order to assess the duration of CD40L mRNA persistence following induction of gene transcription, actinomycin D at 5 mg/ml was added to the cultures one hour after initiation of culture with PMA and ionomycin. It was observed that CD40L mRNA from patients with SLE had a longer half-life compared to CD40L mRNA isolated from PBMC from healthy controls, suggesting that the stability of the CD40L mRNA was prolonged in the SLE cells. These experiments, still in progress, raise several possible interpretations that are being pursued in the context of our funded RO-1 grant "CD40 Ligand Expression in SLE". First, the 3' untranslated segment of the CD40L gene, which confers message stability for many gene products, may be altered in SLE; second, the regulatory proteins that interact with the 3' untranslated segment of CD40L mRNA may be altered in concentration, structure, or function in SLE; or third, and somewhat overlapping with the second possibility, the chronic activation, or an increased propensity for activation, of a broad spectrum of autoantigen reactive T cells in SLE may confer an activation profile, including increased expression or phosphorylation of signaling or regulatory proteins, that promotes prolonged CD40L mRNA expression.

EXAMPLE 3

Analysis of CD40L transcription factor binding motifs

In this Example, the human CD40L proximal promoter was investigated to define the key regulatory factors that mediate gene transcription, and thereby gain insight into the altered regulation of CD40L in systemic autoimmune disease. A previous study of the human CD40L promoter used the Jurkat T-cell line as a source of potential binding proteins (63). In this Example, primary human T-cells were used.

Nuclear extracts from PMA and ionomycin-activated peripheral blood cells were analyzed to examine the proteins from primary human T cells that bind to the human CD40L promoter. The −88 to −57 bp (NM1-L) segment of the promoter contains a classic NF-AT binding motif at the -62 to -69 bp, which location we confirmed the presence of by EMSA. PBMC from a healthy subject were cultured with medium alone, or for 0.5 or 2 hours with PMA and ionomycin, and nuclei were isolated. ^{32}P-labeled double-stranded oligonucleotide fragments, corresponding to −88 to −57 bp (NM1-L) of the proximal CD40L promoter, were incubated with the nuclear extracts and then run on a polyacrylamide gel. Protein complexes bound to the oligonucleotides retarded the migration of the labeled promoter fragments. An oligonucleotide containing a known NF-AT site in the human IL-4 promoter served as a positive control. As demonstrated in **FIG. 1**, protein complexes from activated normal peripheral blood T cells bound to oligonucleotides derived from the proximal CD40L promoter. Binding of that complex was specifically inhibited by pre-incubation of the nuclear extracts with polyclonal anti-NF-AT antibody, but not by incubation with anti-Fos antibody.

In addition to NF-AT, an additional nuclear complex was identified that was present in 2 hour-activated PMA and ionomycin-stimulated cultures, but not in unstimulated T cells. This complex bound to the ^{32}P-labeled oligonucleotide fragment corresponding to bp -73 to −41 (NM1-P) of the proximal CD40L promoter that extends 3' of the proximal NF-AT site, but was not altered by pre-incubation with anti-NF-AT.

These experiments confirmed the capacity of the proximal CD40L promoter to bind at least two nuclear protein complexes, including NF-AT, from activated peripheral blood T cells. The location of the oligonucleotides studied is within the 90 bp just

proximal to the transcription start site of the CD40L promoter and is 3' of the poly-A tract which we have found to be altered in some patient samples (see below).

EXAMPLE 4

Analysis of the functional significance of the altered CD40L promoter

This example describes the transcriptional activity effect of the altered proximal promoter sequence on CD40L gene expression.

The transcriptional activities of wild-type and altered CD40L promoters was compared. To determine whether an alteration of A to C at position -125 of the CD40L proximal promoter affects promoter activity, promoter segments amplified by Pcd1 and Pcd2, with either A or C at the -125 site, were inserted into the luciferase reporter vector pGL-2/basic (Promega). *See* **FIG. 10.** The constructs containing wild-type and altered CD40L promoter fragments were used to transiently transfect human Jurkat T cells by a modified DEAE-dextran electroporation method. 48 hours later, the cells were lysed and assayed for luciferase activity. One hour prior to harvest, an aliquot of the transfected cells was stimulated with PMA 20 ng/ml and ionomycin 500 ng/ml. A construct containing the β-galactosidase gene (Galacto-Light Plus chemiluminescent reporter assay kit – Tropix, Bedford, MA) was co-transfected, and the β-galactosidase activity measured as an internal control to calibrate the transfection efficiency.

The altered CD40L promoter generated a luciferase signal that was 4-fold higher than the wild-type promoter when transfected Jurkat cells were assayed in the absence of stimulation (*See* **FIG. 7**). After activation with PMA and ionomycin for 1 hour, the Jurkat T cells transfected with the promoter expressing the C at position -125 showed 6-fold greater induction of luciferase activity than the wild-type construct. These data suggest that a change from A to C in the poly-A tract of the CD40L proximal promoter confers increased transcriptional activity in a T-cell line system. This experiment can be repeated and extended to transfection of primary T-cells activated with ConA, according to a method described by Cron (109).

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and the accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

It is further to be understood that values are approximate, and are provided for description.

Patents, patent applications, and publications are cited throughout this application, the disclosures of which are incorporated herein by reference in their entireties for all purposes.

BIBLIOGRAPHY

1. Banchereau J, Bazan F, Blanchard F, et al. The CD40 antigen and its ligand. Annu Rev Immunol 12:881-922, 1994.

2. Lederman S, Yellin MJ, Krichevsky A, et al. Identification of a novel surface protein on activated CD4+ T cells that induces contact dependent B cell differentiation. J Exp Med 175:1091-1101, 1992.

3. Spriggs MK, Armitage RJ, Strockbine L, et al. Recombinant human CD40 ligand stimulates B cell proliferation and immunoglobulin E secretion. J Exp Med 176:1543-1550, 1992.

4. Hollenbaugh D, Grosmire I, Kullas C, et al. The human T cell antigen gp39, a member of the TNF gene family, is a ligand for the CD40 receptor: Expression of a soluble form of gp39 with B cell costimulatory activity. EMBO J 11:4314-4321, 1992.

5. Lederman S, Yellin MJ, Inghirami G, Lee JJ, Knowles DM, Chess L. Molecular interactions mediating T-B lymphocyte collaboration in human lymphoid follicles: Roles of T-B activating molecule (5c8 antigen) and CD40 in contact-dependent help. J Immunol 149:3817-3826, 1992.

6. Shapira SK, Vercelli D, Jabara HH, Fu SM, Geha RS. Molecular analysis of the induction of immunoglobulin E synthesis in human B cells by interleukin 4 and engagement of CD40 antigen. J Exp Med 175:289-292, 1992.

7. Splawski, JB, Fu SM, Lipsky PE. Immunoregulatory role of CD40 in human B cell differentiation. J Immunol 150:1276-1285, 1993.

8. Grabstein KH, Maliszexshi CR, Shanebeck K, et al. The regulation of T cell-dependent antibody formation in vitro by CD40 ligand and IL-2. J Immunol 150:3141-3147, 1993.

9. Nonoyama S, Hollenbaugh D, Aruffo A, Ledbetter JA, Ochs HD. B cell activation via CD40 is required for specific antibody production by antigen-stimulated human B cells. J Exp Med 178:1097-1102, 1993.

10. Saeland S, Duvert V, Moreau I, Banchereau J. Human B cell precursors proliferate and express CD23 after CD40 ligation. J Exp Med 178:113-120, 1993.

11. Yellin MJ, Sinning J, Covey LR, et al. T lymphocyte T cell-B cell-activating molecule/CD40-L molecules induce normal B cells or chronic lymphocytic leukemia B cells to express CD80 (B7/BB-1) and enhance their costimulatory activity. J Immunol 153:666-674, 1994.

12. Clark EA, Ledbetter JA. How B and T cells talk to each other. Nature 367:425-428, 1994.

13. Durie FH, Foy TM, Masters SR, Laman JD, Noelle RJ. The role of CD40 in the regulation of humoral and cell-mediated immunity. Immunol Today 15:406-410, 1994.

14. Galy AHM, Spits H. CD40 is functionally expressed on human thymic epithelial cells. J Immunol 149:775-782, 1992.

15. Alderson MR, Armitage RJ, Tough TW, Strockbine L, Fanslow WC, Spriggs MK. CD40 expression by human monocytes: regulation by cytokines and activation of monocytes by the ligand for CD40. J Exp Med 178:669-674, 1993.

16. Yellin MJ, Winikoff S, Fortune SM, Crow MK, Lederman S, Chess L. Ligation of CD40 on fibroblasts induces CD54 upregulation and proliferation. J Leukocyte Biol 58:1-8, 1995.17. Yellin MJ, Brett J, Baum D, et al. Functional interactions of T cells with endothelial cells: The role of CD40L-CD40 mediated signals. J Exp Med 182:1857-1864, 1995.

18. Karmann K, Hughes CCW, Schechner J, Fanslow WC, Pober JS. CD40 on human endothelial cells: inducibility by cytokines and functional regulation of adhesion molecules expression. Proc Natl Acad Sci USA 92:4342-4346, 1995.

19. Mach F, Schonbeck U, Sukhova GK, et al. Functional CD40 ligand is expressed on human vascular endothelial cells, and macrophages: implications for CD40-CD40 ligand signaling in atherosclerosis. Proc Natl Acad Sci USA 94:1931-1936, 1996.

20.Kato T, Hakamada R, Yamane H, Nariuchi H. Induction of IL-12 p40 messenger RNA expression and IL-12 production of macrophages via CD40-CD40 ligand interaction. J Immunol 156:3932-3938, 1996.

21. Fuleihan R, Ramesh N, Loh R, et al. Defective expression of the CD40 ligand in X chromosome-linked immunoglobulin deficiency with normal or elevated IgM. Proc Natl Acad Sci USA 90: 2170-2173, 1993.

22. Allen RC, Armitage RJ, Conley ME, et al. CD40 ligand gene defects responsible for X-linked hyper-IgM syndrome. Science 259:990-993, 1993.

23. Korthauer U, Graf D, Mages HW, et al. Defective expression of T-cell CD40 ligand causes X-linked immunodeficiency with hyper-IgM. Nature 361:539-541, 1993.

24. DeSanto JP, Bonnefoy JY, Gauchat JF, et al. CD40 ligand mutations in X-linked immunodeficiency with hyper-IgM. Nature 361:541-543, 1993.

25. Aruffo A, Farrington M, Hollenbaugh D, et al. The CD40 ligand, gp39, is defective in activated T cells from patients with X-linked hyper-IgM syndrome. Cell 72:291-300, 1993.

26. Ramesh N, Fuleihan R, Ramesh V, et al. Deletions in the ligand for CD40 in X-linked immunoglobulin deficiency with normal or elevated IgM (HIGMX-1). Int Immunol 5:769-773, 1993.

27. Conley ME, Larche M, Bonagura VR, et al. Hyper IgM syndrome associated with defective CD40-mediated B cell activation. J. Clin. Invest. 94:1404-1409, 1994.

28. Koshy M, Berger D, Crow MK. Increased expression of CD40 ligand on systemic lupus erythematosus lymphocytes. J Clin Invest 98:826-837, 1996.

29. Desai-Mehta A, Lu L, Ramsey-Goldman R, Datta SK. Hyperexpression of CD40 ligand by B and T cells in human lupus and its role in pathogenic autoantibody production. J Clin Invest 97:2063 -2073, 1996.

30. Rissoan MC, Van Kooten C, Chomarat P, et al. The functional CD40 antigen of fibroblasts may contribute to the proliferation of rheumatoid synovium. Clin Exp Immunol 106:481-490, 1996.

31. MacDonald KPA, Nishioka Y, Lipsky P, Thomas R. Functional CD40 ligand is expressed by T cells in rheumatoid arthritis. J Clin Invest 9:2404-2414, 1997.

32. Devi BS, Van Noordin S, Krausz T, Davies KA. Peripheral blood lymphocytes in SLE - hyperexpression of CD154 on T and B lymphocytes and increased number of double negative T cells. J Autoimmun 11:471-475, 1998.

33. Wagner UG, Kurtin PJ, Wahner A, et al. The role of CD8+ CD40L+ T cells in the formation of germinal centers in rheumatoid arthritis. J Immunol 161:6390-6397, 1998.

34. Sekine C, Yagita H, Miyasaka N, Okumura K. Expression and function of CD40 in rheumatoid arthritis. J Rheumatol 25:1048-1053, 1998.

35. Harigai M, Hara M, Nakazawa S, et al. Ligation of CD40 induced tumor necrosis factor-alpha in rheumatoid arthritis: a novel mechanism of activation of synoviocytes. J Rheumatol 26:1035-1043, 1999.

36. Grammer AC, Bergman MC, Miura Y, Fujita K, Davis LS, Lipsky PE. The CD40 ligand expressed by human B cells costimulates B cell responses. J Immunol 154:4996-5010, 1995.

37. Miyashita T, McIlraith MJ, Grammer AC, et al. Bidirectional regulation of human B cell responses by CD40-CD40 ligand interactions. J Immunol 158:4620-4633, 1997.

38. Schattner EJ, Mascarenhas J, Reyfman I, et al. Chronic lymphocytic leukemia B cells can express CD40 ligand and demonstrate T-cell type immunostimulatory capacity. Blood 91:2689-2697, 1998.

39. Grammer AC, MacFarland RD, Heaney J, Darnell BF, Lipsky PE. Expression, regulation, and function of B cell-expressed CD154 in germinal centers. J Immunol 163:4150-4159, 1999.

40. Kawabe T, Naka T, Yoshida K, et al. The immune responses in CD40-deficient mice: impaired immunoglobulin class switching and germinal center formation. Immunity 1:167-178, 1994.

41. Waldschmidt TJ, Elsemore J, Noelle RJ, Flavell RA. Mice deficient for the CD40 ligand. Immunity 1:423-431, 1994.

42. Graf D, Muller S, Korthauer U, Van Kooten C, Weise C, Kroczek RA. A soluble form of TRAP (CD40 ligand) is rapidly released after T cell activation. Eur J Immunol 25:1749-1754, 1995.

43. Pietravalle F, Lecoanet-Henchoz S, Blasey H, et al. Human native soluble CD40L is a biologically active trimer, processed inside microsomes. J Biol Chem 271:5965-5967, 1996.

44. Pietravalle F, Lecoanet-Henchoz S, Aubry JP, Elson G, Bonnefoy JY, Gauchat JF. Cleavage of membrane-bound CD40 ligand is not required for inducing B cell proliferation and differentiation. Eur J Immunol 26:725-728, 1996.

45. Burlinson EL, Graber P, Bonnefoy JY, Ozanne BW, Cushley W. Soluble CD40 ligand induces expression of CD25 and CD23 in resting human tonsillar B lymphocytes. Eur J Immunol 26:1069-1073, 1996.

46. Vakkalanka RK, Woo C, Kirou KA, Koshy M, Berger D, Crow MK. The soluble form of CD40 ligand is present in SLE sera and can induce B cell activation. Arthritis Rheum 42:871-881, 1999.

47. Kato K, Santana-Sahagun E, Rassenti LZ, et al. The soluble CD40 ligand sCD154 in systemic lupus erythematosus. J Clin Invest 104:947-955, 1999.

48. Lederman S, Yellin MJ, Cleary AM, et al. T-BAM/CD40-L on helper T lymphocytes augments lymphokine-induced B cell Ig isotype switch recombination and rescues B cells from programmed cell death. J Immunol 152:2163-2171, 1994.

49. Gascan HJ, Gauchat JF, Aversa G, Vlasselaer PV, de Vries JE. Anti-CD40 monoclonal antibodies or CD4+ T cell clones and IL-4 induce IgG4 and IgE switching in purified human B cells via different signaling pathways. J Immunol 147:8-13, 1991.

50. Defrance T, Vanbervliet B, Briere F, Durand I, Rousset F, Banchereau J. Interleukin-10 and transforming growth factor b cooperate to induce anti-CD40-activated naive human B cells to secrete immunoglobulin A. J Exp Med 175:671-682, 1992.

51. Fujieda S, Zhang K, Saxon A. IL-4 plus CD40 monoclonal antibody induces human B cell gamma subclass specific isotype switch: switching to g1, g3, and g4, but not g2. J Immunol 155:2318-2328, 1995.

52. Malisan F, Briere F, Bridon J-M, et al. Interleukin-10 induces immunoglobulin G isotype switch recombination in human CD40-activated naive B lymphocytes. J Exp Med 183:937-947, 1996.

53. Lin SC, Stavnezer J. Activation of NF-kappaB/Rel by CD40 engagement induces the mouse germ line immunoglobulin Cgamma1 promoter. Mol Cell Biol 16:4591-4603, 1996.

54. Cerutti A, Schaffer A, Shah S, et al. CD30 is a CD40-inducible molecule that negatively regulates CD40-mediated immunoglobulin class switching in non-antigen-selected human B cells. Immunity 9:247-256, 1998.

55. Iciek LA, Delphin SA, Stavnezer J. CD40 cross-linking induces Ige germline transcripts in B cells via activation of NF-kB. J Immunol 158:4769-4779, 1997.

56. Cerutti A, Zan H, Schaffer A, et al. CD40 ligand and appropriate cytokines induce switching to IgG, IgA, and IgE and coordinated germinal center and plasmacytoid phenotype differentiation in a human monoclonal IgM+ IgD+ B cell line. J Immunol 160:2145-2157, 1998.

57. Boussiotis VA, Nadler LM, Strominger JL, Goldfeld AE. Tumor necrosis factor is an autocrine growth factor for normal human B cells. Proc Natl Acad Sci USA 91:7007-7011, 1994.

58. Barrett TB, Shu G, Clark EA. CD40 signaling activates CD11a/CD18 (LFA-1)-mediated adhesion in B cells. J Immunol 146:1722-1729, 1991.

59. Schattner EJ, Elkon KB, Yoo D-H, et al. CD40 ligation induces Apo-1/Fas expression on human B lymphocytes and facilitates apoptosis through the Apo-1/Fas pathway. J Exp Med 182:1557-1565, 1995.

60. Schattner EJ, Mascarenhas J, Bishop J, et al. CD4+ T-cell induction of Fas-mediated apoptosis in Burkitt's lymphoma B cells. Blood 88:1375-1382, 1996.

61. Schattner E, Friedman SM. Fas expression and apoptosis in human B cells. Immunol Res 15:246-257, 1996.

62. Ju ST, Panka DJ, Cui H, et al. Fas (CD95)/FasL interactions required for programmed cell death after T-cell activation. Nature 373:444-447, 1995.

63. Schubert LA, King G, Cron RQ, Lewis DB, Aruffo A, Hollenbaugh D. The human gp39 promoter. Two distinct nuclear factors of activated T cell protein-binding elements contribute independently to transcriptional activation. J Biol Chem 270:29624-29627.

64. Perez-Melgosa M, Hollenbaugh D, Wilson CB. CD40 ligand is a limiting factor in the humoral repsonse to T cell-dependent antigens. J Immunol 163:1123-1127, 1999.

65. Allen RC, Spriggs MK, Belmont JW. Dinucleotide repeat polymorphism in the human CD40 ligand gene. Hum Mol Genet 6:828, 1993.

66. Villa A, Notarangelo LD, Di Santo JP, et al. Organization of the human CD40L gene: implications for molecular defects in X chromosome-linked hyper-IgM syndrome and prenatal diagnosis. Proc Natl Acad Sci USA 91:2110-2114, 1994.

67. Shimadzu M, Nunoi H, Terasaki H, et al. Structural organization of the gene for CD40 ligand: molecular analysis for diagnosis of X-linked hyper-IgM syndrome. Biochim Biophys Acta 1260:67-72, 1995.

68. Seyema K, Kira S, Ishidoh K, Souma S, Miyakawa T, Kominami E. Genomic structure and PCR-SSCP analysis of the human CD40 ligand gene: its application to prenatal screening for X-linked hyper-IgM syndrome. Hum Genet 97:180-185, 1996.

69. Tsytsykova AV, Tsitsikov EN, Geha RS. The CD40L promoter contains nuclear factor of activated T cells-binding motifs which require AP-1 binding for activation of transcription. J Biol Chem 271:3763-3770, 1996.

70. Cron RQ, Schubert LA, Lewis DB. Identification of an enhancer element located downstream of the CD40-ligand gene. Arthritis Rheum 41:S235, 1998.

71. Mohan C, Shi Y, Laman JD, Datta SK. Interaction between CD40 and its ligand gp39 in the development of murine lupus nephritis. J Immunol 154:1470-1480, 1995.

72. Kirk AD, Burkly LC, Batty DS, et al. Treatment with humanized monoclonal antibody against CD154 prevents acute renal allograft rejection in nonhuman primates. NatureMed 5:686-693, 1999.

73. Webster EA, Khakoo AY, Mackus WJ, et al. An aggressive form of polyarticular arthritis in a man with CD154 mutation (X-linked hyper-IgM syndrome). Arthritis Rheum 42:1291-1296, 1999.

74. Durie FH, Fava RA, Foy TM, Aruffo A, Ledbetter JA, Noelle RJ. Prevention of collagen-induced arthritis with an antibody to gp39, the ligand for CD40. Science 261:1328-1330, 1993.

75. Gomolka M, Menninger H, Saal JE, et al. Immunoprinting: various genes are associated with increased risk to develop rheumatoid arthritis in different groups of adult patients. J Mol Med 73:19-29, 1995.

76. Shiozawa S, Hayashi S, Tsukamoto Y, et al. Identification of the gene loci that predispose to rheumatoid arthritis. Int Immunol 10:1891-1895, 1998.

77. Storb U. The molecular basis of somatic hypermutation of immunoglobulin genes. Curr Opin Immunol 8:206-214, 1996.

78. Parsons R, Myeroff LL, Liu B, Willson JK, et al. Microsatellite instability and mutations of the transforming growth factor beta type II receptor gene in colorectal cancer. Cancer Res 55:5548-5550, 1995.

79. Markowitz S, Wang J, Myeroff L, et al. Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. Science 268:1336-1338, 1995.

80. Renault B, Calistri D, Buonsanti G, Nanni O, Amadori D, Ranzani GN. Microsatellite instability and mutations of p53 and TGF-beta RII genes in gastric cancer. Hum Genet 98:601-607, 1996.

81. Huang J, Papadopoulos N, McKinley AJ, et al. APC mutations in colorectal tumors with mismatch repair deficiency. Proc Natl Acad Sci USA 93:9049-9054, 1996.

82. Peng H, Chén G, Du M, Singh N, Issacson PG, Pan L. Replication error phenotype and p53 gene mutation in lymphomas of mucosa-associated lymphoid tissue. Am J Pathol 148:643-648, 1996.

83. Akiyama Y, Tsubouchi N, Yuasa Y. Frequent somatic mutations of hMSH3 with reference to microsatellite instability in hereditary nonpolyposis colorectal cancers. Biochem Biophys Res Commun 236:248-252, 1997.

84. Takenoshita S, Tani M, Nagashima M, et al. Mutation analysis of coding sequences of the entire transforming growth factor beta type II receptor gene in sporadic human colon cancer using genomic DNA and intron primers. Oncogene 14:1255-1258, 1997.

85. Laken SJ, Petersen GM, Gruber SB, et al. Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC. Nature Gen 17:79-83, 1997.

86. Venkatasubbarao K, Ahmed MM, Swiderski C, et al. Novel mutations in the polyadenine tract of the transforming growth factor beta type II receptor gene are found in a subpopulation of human pancreatic adenocarcinomas. Genes Chrom Cancer 22:138-144, 1998.

87. Guo RJ, Wang Y, Kaneko E, et al. Analyses of mutation and loss of heterozygosity of coding sequences of the entire transforming growth factor beta type II receptor gene in sporadic human gastric cancer. Carcinogenesis 19:1539-1544, 1998.

88. Tran HT, Keen JD, Kricker M, Resnick MA, Gordenin DA. Hypermutatability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants. Mol Cell Biol 17:2859-2865, 1997.

89. Sia EA, Kokoska RJ, Dominska M, Greenwell P, Petes TD. Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes. Mol Cell Biol 17:2851-2858, 1997.

90. Vidaud D, Vidaud M, Bahnak BR, et al. Haemophilia B due to a de novo insertion of a human-specific Alu subfamily member within the coding region of the factor IX gene. Eur J Hum Genet 1:30-36, 1993.

91. Ellinger T, Behnke D, Knaus R, Bujard H, Gralla JD. Context-dependent effects of upstream A-tracts. Stimulation or inhibition of Escherichia coli promoter function. J Mol Biol 239:466-475, 1994.

92. Aiyar SH, Gourse RL, Ross W. Upstream A-tracts increase bacterial promoter activity through interactions with the RNA polymerase a subunit. Proc Natl Acad Sci USA 95:14652-14657, 1998.

93. Shain DH, Zuber MX, Norris J, Yoo J, Neuman T. Selective conservation of an E-protein gene promoter during vertebrate evolution. FEBS Lett 440:332-336, 1998.

94. Shain DH, Zuber MX, Neuman T. Transcription initiation from a poly(dA) tract. Nucleic Acids Res 26:1019-1025, 1998.

95. Tang Y, Nilsson L. Interaction of human SRY protein with DNA: a molecular dynamics study. Proteins 31:417-433, 1998.

96. Magistrelli C, Gergely P Jr, Samoilva E, et al. Polymorphic genotypes of the HRES-1 human endogenous retrovirus locus are associated with SLE. Arthritis Rheum 42:S386, 1999.

97. Parekh BS, Maniatis T. Virus infection leads to localized hyperacetylation of histones H3 and H4 at the IFN-beta promoter. Mol Cell 3:125-129, 1999.

98. Sherka C, Decker EL, Zipfel PF. Coordinate expression and distinct DNA-binding characteristics of the four EGR-zinc finger proteins in Jurkat T lymphocytes. Immunobiology 198:179-191, 1997.

99. Hodge MR, Ranger AM, Charles de la Brousse F, Hoey T, Grusby MJ, Glimcher LH. Hyperproliferation and dysregulation of IL-4 expression in NF-ATp-deficient mice. Immunity 4:397-405, 1996.

100. Newton CR, Graham A, Heptinstall LE, Powell SJ, Summers C, Kalsheker N, et al. Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). Nucleic Acids Res 17:2503-2516, 1989.

101. Li Y, Sun G-R, Zheng Q, et al. Allelic variants of human TCR BV17S1 defined by restriction fragment length polymorphism, single strand conformation polymorphism, and amplification refractory mutation system analysis. Hum Immunol 49:85-95, 1996.

102. Quandt K, Frech K, Karas H, Wingender E, Werner T. MatInd and MatInspector – new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res 23:4878-4884, 1995.

103. Ellegren H. Variable SINE 3' poly(A) sequences: an abundant class of genetic markers in the pig genome. Mamm Genome 4:429-434, 1993.

104. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77:51-59, 1989.

105. Takashiba S, Shapira L, Amar S, Van Dyke TE. Cloning and characterization of human TNFa promoter region. Gene 131:307-308, 1993.

106. Perfetto C, Zacheis M, McDaid D, Meador JW 3rd, Schwartz BD. Polymorphism in the promoter region of HLA-DRB genes. Hum Immunol 36:27-33, 1993.

107. Deng X, Sun G-R, Zheng Q, Li Y. Characterization of human TCR Vb gene promoter. Role of the dodecamer motif in promoter activity. J Biol Chem 273:23709-23715, 1998.

108. Gonzalez CM, Cristobal E, Bascones M, et al. Collagen type I alpha polymorphism at Sp1 binding site and bone mineral density. Arthritis Rheum 42:S232, 1999.

109. Cron RQ, Schubert LA, Lewis DB, Hughes CC. Consistent transient transfection of DNA into non-transformed human and murine T-lymphocytes. J Immunol Methods 205:145-150, 1997.

110. Lai CK, Ting LP. Transcriptional repression of human hepatitis B virus genes by a bZIP family member, E4BP4. J Virol 73:3197-3209, 1999.

111. Akira S, Isshiki H, Sugita T, et al. A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. EMBO J 9:1897-1906, 1990.

112. Pope RM, Leutz A, Ness SA. C/EBP regulation of the tumor necrosis factor a gene. J Clin Invest 94:1449-1455, 1994.

113. Benbrook DM, Jones NC. Different binding specificities and transactivation of variant CRE's by CREB complexes. Nucleic Acids Res 22:1463-1469, 1994.

114. Greene CN, Jinks-Robertson S. Frameshift intermediates in homopolymer runs are removed efficiently by yeast mismatch repair proteins. Mol Cell Biol 17:2844-2850, 1997.

115. Tuffery S, Moine P, Demaille J, Claustres M. Identification of variable length polyadenosine tract at the dystrophin locus. Hum Genet 95:590-592, 1995.

116. Ausubel *et al.* (eds.), Current Protocols in Molecular Biology, John Wiley & Sons, Inc.,1994.

117. Freshney RI (ed), Animal Cell Culture, 1986.

118. Glover (ed.), DNA Cloning: A Practical Approach, Volumes I and II, 1985.

119. Hames & Higgins (eds.), Transcription And Translation, 1984.

120. Hames & Higgins (eds.), Nucleic Acid Hybridization, 1985.

121. Perbal, A Practical Guide To Molecular Cloning, 1984.

122. Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

123. Vakkalanka RK, Kirou KA, Raman C, Crow MK: Kinetics of CD40 ligand (CD40L) mRNA expression in systemic lupus erythematosus (SLE) patients and healthy controls. Proceedings of the 10th International Congress of Immunology, 1998.

124. Kirou KA, Crow MK: New pieces to the SLE cytokine puzzle. Clin Immunol, 91:1-5, 1999.

125. Crow MK: Mechanisms of T-helper cell activation and function in systemic lupus erythematosus. In: Lupus: Molecular and Cellular Pathogenesis. Edited by G Kammer and G Tsokos, Totowa, NJ: Humana Press, Inc., pp. 231-256, 1999.

126. Li Y, Sun G-R, Crow MK: An altered nucleotide sequence in the immediate promoter region of CD40 ligand is associated with rheumatoid arthritis. Arthritis Rheum 42:S233, 1999.

127. Köhler and Milstein, Nature 1975;256:495-497.

128. Kozbor *et al.*, Immunology Today 1983;4:72 *et seq.*

129. Cole *et al.*, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96, 1985.

130. Cote *et al.*, Proc. Natl. Acad. Sci. U.S.A. 1983;80:2026-2030.